Interactions of Technology and Design in Nanoscale SRAM

R.W. Mann

Department of Electrical and Computer Engineering
University of Virginia

Dissertation Proposal Presented in Partial Fulfillment of the
Requirement for the
Doctor of Philosophy Degree
in Electrical Engineering

December 10, 2009

Interactions of Technology and Design in Nanoscale SRAM

Abstract

Continued advances in silicon technology have enabled the VLSI industry to shrink the area of the transistor by roughly a factor of two with each successive technology node. This trend has continued unabated for five decades and has made computing a ubiquitous entity in modern culture. Made possible by continuous advances in CMOS technology and fueled by a growing and fiercely competitive market, in order to continue through the next decade will require continued advances in CMOS process technology as well as circuit design innovation.

The 6T SRAM cell design has been successfully scaled in both bulk and SOI down to the 32/28nm node and has remained for more than a decade the dominant technology development vehicle for advanced CMOS technologies. Reduced device dimensions and operating voltages that accompany technology scaling have led to increased design challenges with each successive technology node. Thus, reduced functional yield margins coupled with increasing variability of the CMOS device characteristics have become the most significant problem facing future nanoscale SRAM, motivating this effort.

To address these challenges, the proposed approach first develops a custom scaled (90nm-22nm) predictive technology model (PTM) based framework, calibrated using published industry target values to quantify and address the challenges confronting nanoscale SRAM below the 65nm node. The role and contribution of the 6T cell design topology in addressing variation is proposed using technology computer aided design (TCAD) tools. While circuit assist methods have shown promise in extending the life of the 6T SRAM, this work proposes to quantify and also develop a sensitivity based methodology for assessing the effectiveness of the assist methods in addressing the reduced functional margins. Additionally, a new margin/delay analysis is proposed as a means of assessing the functional effectiveness of the circuit assist methods. The margin/delay analysis may be further extended to assess the limits of circuit assist methods in extending the 6T SRAM beyond 32nm node. Finally, a constraint based analysis is proposed to assess the extent to which these methods may provide effective solutions as the technologies are scaled beyond the 32/28nm node.

CHAPTER 1

Introduction – SRAM and CMOS Technology Scaling

-Abstract

The commercial success and ubiquitous computing power ranging from handheld and portable devices to mainframe supercomputers has been made possible by the reduced cost per memory bit and logic gate with each successive technology generation. This reduced cost is made possible by continued advances in CMOS device scaling. The design challenges such as increased variability and standby power which result from scaling through the nanoscale regime are even more acutely exhibited in the dense SRAM devices which employ sub-minimum design rules. SRAM remains the most cost effective embedded memory solution for many applications; however, fundamental challenges arise as technologies continue to scale below 100nm. This chapter defines the rapidly emerging challenges facing CMOS SRAM technologies in the nanoscale era and defines the problem set to be addressed and the scope of the work.

Introduction

As scaling continues, the increasing variability that accompanies reduced dimensions coupled with reduced overdrive associated with lower operating voltages has created significant challenges beyond 65nm. This is especially true for the 6T SRAM cell that often uses minimum device dimensions and requires many sub-minimum design rules to achieve the aggressive density targets. The 6T SRAM cell design has been successfully scaled across many technology generations and because it generally requires little deviation from base logic processing, is frequently used as the technology development vehicle for advanced CMOS technologies. Reduced device dimensions and operating voltages that accompany technology scaling have led to increased design challenges with each successive technology node.

Despite these challenges, the 6T SRAM is expected to continue to play a dominant role in future technology generations because of its combination of density, performance, and compatibility with logic processing. The successful commercial scaling of the 6T SRAM driven by strong industry competition has followed a well defined linear shrink factor of 0.7X over multiple generations resulting in a predictable 2X reduction in cell area per generation. Despite the numerous challenges in lithography, device, and process integration the trend in 6T bit cell area is expected to continue beyond the 32nm node. This trend in 6T cell area, shown in Fig. 1, is projected out to the 22nm generation where the competitive 6T cell size is expected to be approximately $0.0625\mu m^2$. This continued trend in area reduction is accompanied by the well known consequence of increased variability associated with the reduced channel area. Although technology options such as high-κ with metal gate have provided some relief in variability, the level of integration and functional margins beyond the 28/32nm generation pose a serious technical challenge.

To investigate the impact of scaling, customized LP PTMs for nodes from 90nm to 22nm for a conventional poly silicon/nitrided-silicon dioxide stack were developed that were calibrated with published industry LP CMOS data [1] [2] [3] [4] . Table I provides the essential metrics used and technology scaling assumptions for the LP models in our study.

Table 1-1 Custom low power technology models 65nm-22nm

	Node	**65**		**45**		**32**		**22**	
	device type	N	P	N	P	N	P	N	P
Target Values	**Vnom (V)**	1.2		1.1		1.1		1	
	Tox (nm)	2		1.8		1.6		1.4	
	Lpoly (nm)	56		39		27		19	
	Ion (uA/um)	600	300	620	300	700	380	720	380
	Ioff (pA/um)	250		400		1000		2000	
	HVT Ion (uA/um)	400	210	410	210	440	340	450	340
	HVT Ioff (pA/um)	10		30		50		150	
Tuned model values	**Ion (uA/um)**	606	305	615	309	709	381	725	385
	Ioff (pA/um)	250	219	477	409	947	965	1858	1915
	HVT Ion (uA/um)	409	220	425	229	444	330	469	331
	HVT Ioff (pA/um)	10	9	36	35	45	62	183	172

Because variation is both an intrinsic property of the technology and a fundamental limiter to scaling, the most significant contributors of variation were incorporated into the models. Threshold voltage (V_T) variation due to random dopant fluctuations (RDF) is proportional to $1/\sqrt{W \cdot L}$ as described by Pelgrom et al [5]. The where the mismatch between two identically defined transistors in close proximity to one another may be described by:

$$\sigma\Delta V_T = A_{Vt} \cdot {}^{1}\!\Big/\!_{\sqrt{W \cdot L}} \tag{1-1}$$

where the quantity A_{Vt} has units of $mV - \mu m$ and W and L refer to the device width and length respectively. What has now become commonly referred to as the Pelgrom plot, where the delta V_T of two identically drawn, adjacent devices is an important analysis tool in SRAM cell development to enable separation of the systematic from the intrinsic variation sources. This fact will be discussed further in chapter 3. The value for A_{Vt} used was based on published hardware measurements [6] [7] for competitive industry technologies. The channel length variation (both global and local), and the variation in V_T associated with implant dose variations were also included. The combined effects of scaled gate oxide thickness (Tox) of approximately 10% per generation over the range of technologies included and corresponding increase in the effective channel doping (N) of approximately 20% per generation tend to hold the A_{Vt} values roughly constant with each generation, which may be explained by the commonly used empirical equation [8]:

$$\sigma\Delta V_T = 3.19 \cdot 10^{-8} \cdot T_{ox} \cdot \frac{\mathrm{N}^{0.4}}{\sqrt{W \cdot L}} [V] \tag{1-2}$$

For clarity, it should be noted that the $\sigma\Delta V_T$ refers to the mismatch between two identically defined devices in close proximity while the V_T σ for an individual device is therefore smaller by a factor of $1/\sqrt{2}$. The implications of the scaled devices employed in the SRAM cell are shown in Figure 2. Both components assumed a 3sigma value equal to 10% of the target ($L_{physical}$) for the technology. A 30mV (3sigma) variation in V_{TO} for NMOS and PMOS due to implant dose and energy variability was used.

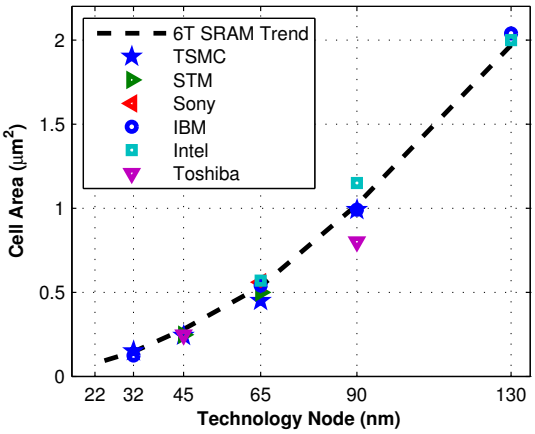


Figure 1-1 Trend in SRAM cell size with scaling based on published cell sizes.

The infrastructure resulting from this analysis coupled with the known scaling relationship for the SRAM devices from Fig. 1 provides a means of assessing the local variation in threshold voltage across the technology nodes. This is relationship is shown in Fig. 2 where the local variation in SRAM device threshold voltage has increased for the 32nm node by roughly a factor of 2 over the variation addressed at the 90nm node. The adoption of high-κ with metal gate beyond 45nm will provide some relief, consistent with (1-2), but the increasing trend will again increase as the channel area is reduced through scaling.

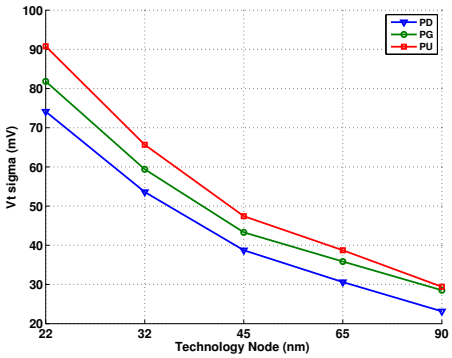


Figure 1-2 Impact of scaling trends on pull down (PD), pass gate (PG) and pull up (PU) SRAM device Vt sigma based on the RDF component.

In addition to the increased variation, voltage scaling has several additional consequences for the SRAM design. The functional margins, read noise margin and write margin are reduced as shown in Fig. 3. The read static noise margin (RSNM) or (SNM) is a measure of the stability of the cell during access [9]. The ability of the cell to retain

state is a fundamental requirement for the successful operation of the SRAM. This is not only true during a read operation but also for the cells in unselected columns on an asserted word line. This is often referred to as the half-select issue. The write margin (WM) is a measure of the ability to write data to the cell. This margin is also trending to downward with scaling.

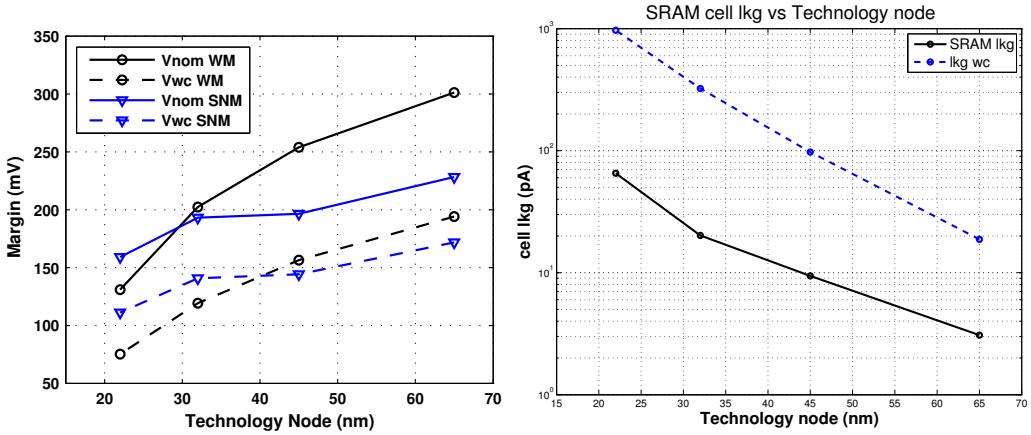


Figure 1-3 Functional and parasitic leakage trends in 6T SRAM with continued technology scaling. (a) Functional margins are decreasing (both WM and RSNM) with scaling and (b) parasitic leakage from source current (Ioff), gate tunneling current and gate induced drain leakage (GIDL) are increasing.

While the functional margins are decreasing with continued scaling, the standby power for the array, as measured by the bit cell parasitic leakage, is increasing. There are three primary mechanisms involved in this trend. First the gate tunneling current is increasing with Tox reduction. The tunneling mechanisms are voltage accelerated and exhibit little temperature dependence. The second is gate induced drain leakage (GIDL). The use of halo or pocket implants to improve the short channel effects (SCE) by reducing drain induced barrier lowering (DIBL) has tended to increase GIDL in the devices. For low power technologies, GIDL may be a significant component of the off state parasitic leakage. The third major leakage contributor is sub-threshold leakage. This mechanism is governed by the sub-threshold slope and the threshold voltage of the device by the following relationship:

$$I_{sub} = I_0 \cdot \frac{W}{L} \cdot 10^{\left(\frac{Vgs-Vt}{S}\right)} \tag{1-3}$$

where I_0 is a technology dependent constant with units of current, Vgs is the gate to source voltage (0V in off state), V_T is the threshold voltage and S is the sub-threshold slope. S has units of mV/decade and is expressed in the relationship [10] (1-4).

$$S = \ln(10) \cdot \frac{k \cdot T}{q} \cdot \left(1 + \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \cdot \frac{Tox}{X_d}\right) \cdot \left(1 + \frac{11 \cdot Tox}{X_d} \cdot \exp\left(\frac{-\pi \cdot Leff}{2 \cdot X_d + 3 \cdot Tox}\right)\right) \tag{1-4}$$

In expression (1-4), Tox is the gate oxide thickness, k is Boltzman's constant, T is temperature in Kelvin, q is the fundamental charge of an electron in eV, X_d is the

depletion thickness, Leff is the effective channel length. The threshold voltage also tends to decrease with the technology Vdd in order to achieve sufficient overdrive to preserve performance. Because GIDL and gate leakage are tunneling mechanisms and exhibit little temperature dependence, the sub-threshold leakage becomes the dominant leakage source at elevated temperatures. The technology choice is often a critical factor in the array standby power. Technology solutions optimized for low SRAM standby power have achieved leakage values averaging < 50fA/cell at 25C [11].

As can be observed from (1-3) and (1-4), the introduction of high–κ dielectric materials for the gate dielectric can improve the sub-threshold slope by allowing reduced Tox values and therefore the sub-threshold leakage. The net contribution of the high-k material therefore is significant in providing a path to improve variability, gate leakage and to some extent, sub-threshold leakage. While this is one example demonstrating the close interaction and role of CMOS technology in the future of nanoscale SRAM, many such interactions exist.

Another significant area of concern for scaled 6T SRAM is the susceptibility to radiation induced soft errors in the form of both single event upsets (SEU) and SRAM array multi-bit (MB) fails [12] [13] [14]. The two primary sources of soft error inducing radiation are from either terrestrial radiation or from radioactive isotopes within materials used in the integrated circuit fabrication process. High energy cosmic radiation interacting with the earth's atmosphere results in a flux of neutron particles with a large range of energies extending to several 100MeV. At sea level the resulting high energy neutrons manifest a relatively isotropic flux of ~10-20 neutrons/cm^2 -hr and can interact with the silicon lattice through elastic and inelastic recoil or by spallation where the silicon atom is shattered into heavy and one or several lighter particles. This process produces a charge cloud of electron-hole pairs that, when in close proximity to one or more sensitive neighboring circuit nodes, may result in a single or multi-bit error.

The second form of radiation which predominately originates from impurities within the materials used in modern interconnect technology is the alpha particle. The alpha particle can be characterized as a doubly ionized (He) atom consisting of 2 protons and 2 neutrons. The alpha particle originating from impurities found in the interconnect or packaging materials used in integrated circuit manufacturing has an initial energy extending up to ~8 MeV depending on the specific impurity isotope present. Current purity levels in VLSI processing are sufficient to insure that the alpha flux is not greater than 0.001 α/cm^2-hr. Because the alpha particle is ionized, it interacts with the silicon lattice to produce a column of electron-hole pairs along the path of the particle, which can cause an upset if the charge collected at the circuit node exceeds the critical charge (Q_{crit}) for that circuit or memory bit.

The soft error rate (SER) is expressed as:

$$SER \propto F \times A_{diff} \times \exp\left(-\frac{Q_{crit}}{Q_s}\right) \tag{1-5}$$

where F is the particle flux, A_{diff} is the critical or sensitive charge collection area, Q_{crit} is the critical amount of charge required to flip the bit and Q_s is the charge collection efficiency. Cell designs topologies which minimize A_{diff} and increase Q_{crit} are therefore

preferred. The charge collection efficiency Q_s is modulated by factors such as voltage, charge sharing, NWELL and PWELL depth, use of retrograde well doping profiles and use of triple well. The amount of charge collected at a given node is typically much less than the total charge generated. Values for Qs are obtained following the trends provided by Hazucha and Svensson [15]. Fig. 4 provides a summary of published Qcrit values as well as simulated Qcrit for the scaled technologies defined in this work down to 22nm.

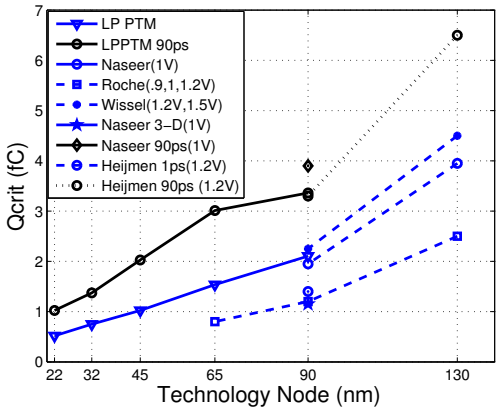


Figure 1-4 Trend in 6T SRAM Qcrit values with continued scaling.

An additional challenge that confronts the nanoscale SRAM design is the shift in threshold voltage during product lifetime. The most significant mechanism for this has been negative bias temperature instability (NBTI) which results in a degradation of the PMOS device associated with a shift in the threshold voltage [16] [17]. This produces in a corresponding shift in the functional margins of the SRAM cell discussed earlier. The SNM will be decreased by an amount typically on the same order of the mean V_T shift of the PMOS device while the write margin will be improved by the weakened PMOS.

The obstacles which confront the continued scaling of the 6T SRAM are formidable and yet for the past five decades the trend has been maintained with limited deviation. One of the most fundamental limits, gate tunneling associated with reduced Tox, has been at least temporarily shifted or postponed by the introduction of high-κ materials. This change has also provided relief for variability and should also improve the sub-threshold slope. The impact on NBTI and PBTI are not yet clear.

Specific Contributions in the field of CMOS SRAM technology to include:

- ***Development of Low Power PTM Models (90-22nm)***
- ***Spice simulation of Qcrit for 90-22nm LP and HP bulk CMOS***
- ***Defining the challenges for SRAM in Nano-scale CMOS process technologies***
- ***Concise description of silicon mechanisms and how they impact SRAM design***

Publications:

a. ***A half micron CMOS logic generation, IBM J. Res. Dev. (USA) Vol.39, No.1-2 Jan. March 1995, p.215-27***
b. ***Silicides and local interconnects for high-performance VLSI applications: a review, IBM Journal of Res. and Dev., Vol.39, No.4 July P403-17 (1995).***
c. ***0.13μm high performance SOI technology development, VLSI Tech Symp paper 19.1 (2000).***
d. ***Enhanced Mult-Threshold (MTCMOS) Circuits Using Variable Well Bias, Low Power Electronics and Design, International Symposium on, 2001. 6-7 Aug. 2001 Page(s):165 - 169***
e. ***High Performance and Low Power Transistors Integrated in 65nm Bulk CMOS Technology, Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International 13-15 Dec. 2004 Page(s):661 - 664***
f. ***65nm CMOS technology for low power applications, Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International 5-7 Dec. 2005 Page(s):64 – 67***
g. ***Ultralow-Power SRAM technology, IBM J. Res. & Dev. Vol. 47, no. 5/6 Sep/Nov 2003, p. 553-566***

Publication @ UVA:
Benton H. Calhoun, Sudhanshu Khanna, Randy Mann, and Jiajing Wang,"Sub-threshold Circuit Design with Shrinking CMOS Devices", International Symposium on Circuits and Systems, 2009.

Related Patents:

- *7,087,486 Method for scalable, low-cost polysilicon capacitor in a planar DRAM*
- *7,057,180 Detector for alpha particle or cosmic ray*
- *6,489,223 Angled implant process*
- *6,187,679 Low temperature formation of low resistivity titanium silicide*
- *7,005,334 Zero threshold voltage pFET and method of making same*
- *6,144,086 Structure for improved latch-up using dual depth STI with impurity implant*
- *6,962,838 High mobility transistors in SOI and method for forming*
- *6,946,376 Symmetric device with contacts self aligned to gate*

CHAPTER 2

SRAM bit cell design for Nanoscale CMOS

-Abstract

The SRAM cell area has become a benchmark of technology competitiveness in today's VLSI industry. The design trade-offs to achieve the aggressive SRAM bit cells are becoming more challenging with each successive technology generation. To achieve the density, performance, and functional requirements, the competitive bit cell requires design rules which are much more aggressive than those used in base logic designs. For this reason the bit cell has become an integral part of the technology offering for the technology suppliers. Because SRAM is largely compatible with CMOS logic processing, and failures can be readily identified through bit fail mapping, it is commonly used by industry as a technology qualification vehicle. Although many design rules are limited directly by lithography, there are several mechanisms which must be addressed for the commercially successful nano-scale cell design.

2.1 Introduction

The complex set of decisions that must be addressed in defining the competitive SRAM bit cell design require a combined understanding of device physics, process integration capabilities, as well as an understanding of the circuit and memory architecture design. Additionally, commercial success will require an understanding of the competitive market as well in order to optimally balance the density, performance, functional margins and power requirements. In this chapter, the interaction of process integration technology in SRAM cell design is explored using technology computer aided design (TCAD) process simulation tools. The Synopsys simulation tools and simulation environment will be used to specifically examine the scaling limitations and challenges for the SRAM cell design.

While the SRAM devices are fundamentally the same, and built concurrently using the same processing steps as the devices used in logic circuits, often significant differences exist. This is because of several factors. First, there may be intended deltas due to the use of additional V_T tailor steps to fine tune the SRAM threshold voltage for optimal functional margin (yield), performance or leakage optimization reasons. The second reason, which is non-intentional, is due to the process, structural differences, stress and proximity effects. For these reasons, commercial nanoscale CMOS technology suppliers provide a set of unique models for the SRAM cell devices to accompany the supplied bit cell.

2.2 Cell Topology

The choice of cell topology is perhaps the most critical choice and must be made early in the technology development phase. This choice will significantly impact the ultimate cell size and aspect ratio that can be obtained as well as compatibility with assist methods. It influences the bit line capacitance and the design rules that will need to be

pushed and if any unique (non-logic based) features will be desired such as the shared contact used in the majority of today's 6T bit cells. Included in the topology decision are the number of transistors to be used and the alternate bit cell options. The term alternate bit cells is used to describe a range of bit cell options that include the total transistor options of 4 through 10 (excluding 6T).

2.2.1 Cell topology: non-6T cells (alternate bit cell options)

With the recognition that achieving the performance, yield and leakage targets with the 6T cell is becoming increasing more difficult to with each technology generation, alternative (non-6T) cell topologies have been proposed. The alternate cells tend to provide a solution that addresses one or more of the challenges highlighted in chapter 1. A few examples of the alternative cells are: 5T which offers a path to improve stability but requires a write assist [18], the 7T [19], there are several implementations of 9T [20] [21] and 10T [22] [23] [24].

Although the alternatives do tend to provide partial solutions, the 8T cell topology [25] is becoming more commonly used in commercial applications, particularly for L1 and L2 cache applications. While the 8T cell area is larger than the 6T, it does offer several advantages. With the two added transistors as a read buffer, the read disturb mechanism is avoided (see further discussion in chapter 3). Because this design is still subject to the half-select disturb during a write operation, array architecture changes are used to avoid this mechanism. Further, this design offers both read and write word lines so additional performance gains can be realized by optimizing the read and write paths independently [26].

2.2.2 6T Cell topologies

As previously mentioned, the optimal 6T topology will be dependent on many factors. These include processing capability, performance, density, power, and functional requirements. There are, at least initially, a number of options theoretically available for placing 6 transistors to perform the desired function. A summary provided by Ishida is shown in Fig. 2-1 for this discussion [27]. Following the nomenclature of Ishida, although published examples of type 2 and 3 can be found, the cell type 1a was the dominant industry topology across several nodes prior to 90nm. Below 90nm the type 4 cell topology became predominantly used in the industry. There are several reasons for this and several important consequences.

As scaling continued below 100nm, the lithographic challenges in printing and controlling the dimensions in orthogonal directions became increasingly difficult. This limitation would also restrict logic layout as well and has an influence on several design rules. The active single crystal regions are printed in one direction and likewise all the polysilicon gates are printed in one direction in this cell type. Of the cell topologies or types summarized by Ishida, only type 4 and a slight variation of type 1-b provide this advantage. Largely as a result of this advantage and the fact that the bit lines would be shorter, crossing the narrow dimension thus lowering the overall capacitance, the type 4 topology became the predominant industry bit cell design beyond 90nm. Because this design style is significantly more area efficient when a shared contact is used to provide the cross couple, this new SRAM feature was adopted. This feature is typically allowed

only in the controlled environment of the SRAM cell and therefore adds a degree of complexity above that of a logic only process.

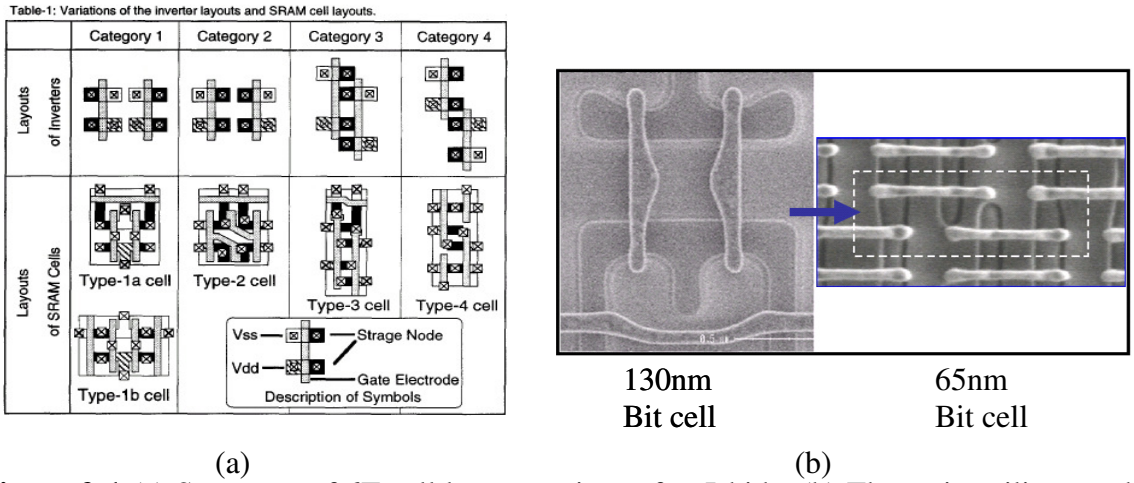


Figure 2-1 (a) Summary of 6T cell layout options after Ishida. (b) The active silicon and poly silicon printed layers for the Intel 130nm [28] and 65nm [29] bit cell designs.

2.2.3. Type 4 Cell Topology

Because the type 4 topology has emerged as the dominant cell design in the industry and has been successfully migrated across 4 technology nodes (90nm-32nm) in both bulk and SOI based technologies, a more careful investigation of this cell is warranted. In addition to added processing complexity associated with the shared contact for this type of cell topology there are two additional consequences that will be addressed. The type 4 topology has a different symmetry from the type1 which have the following ramifications. First, one of the most obvious differences is that the n+ to p+ space appears twice in the direction parallel with the polysilicon patterns. This fact increases the need to push this space much harder for the type 4 topology to achieve the density requirements. Second, the pair of pull down and pass gate NMOS devices share separate active silicon islands and when triple well isolation is used are in separate wells. This is in contrast with the topology of type 1-a. Because the matching of the pull down NMOS devices is a critical factor in the margin analysis, any factor which has the potential to elevate the mismatch between these NMOS devices in the cell is a concern for nanoscale topologies.

The competitive dense bit cells with the type 4 topology have an n+ to p+ space that is roughly 2x tighter than the logic based rules. There are several concerns raised with this pushed rule. TCAD simulation was used to investigate the physical dimensions, and doping profiles representative of the 45nm node. As shown in Fig 2-2, the alignment requirements for the wells must be extremely aggressive in order to prevent implant straggle induced counter doping effects.

2.2.4 Sources of within cell mismatch

An important consideration in the bit cell design is the potential for alignment related mismatch sources. This can arise from several factors and the type 4 topology is particularly vulnerable to this issue. Four such alignment driven sources that can

introduce non-random sources of mismatch are: (1) lateral implanted ion scattering, (2) polysilicon inter-diffusion driven counter-doping, (3) lateral ion straggle from the photo-resist and (4) photo-resist implant shadowing. Of these 4 mechanisms, 1 and 3 are unique to bulk CMOS while 2 and 4 are issues in both bulk and SOI technologies.

The results of lateral ion scattering will exhibit alignment sensitivity and may manifest as mismatch or depending on the well profile a potentially more serious source to drain punch-through path as illustrated in Fig 2-2. Using dimensions and implant profiles consistent with 45nm designs an NWELL mask misalignment of 30nm was sufficient to create a substantial counter-doping path between the source and drain of the PD NMOS device. The well profiles must be carefully tailored and the well alignment tolerances must be adequate to prevent this physical mechanism from occurring.

Polysilicon inter-diffusion is also significant concern with scaling as n+ to p+ space is aggressively pushed. The use of poly pre-doping steps is commonly used to insure the n+ polysilicon is adequately degenerately doped. The alignment of this pre-dope mask as well as the n+ and p+ source drain implant masks must be carefully balanced to avoid the scenario shown in Fig. 2-3. Migration to metal gate and high-κ will avoid this concern beyond 32nm.

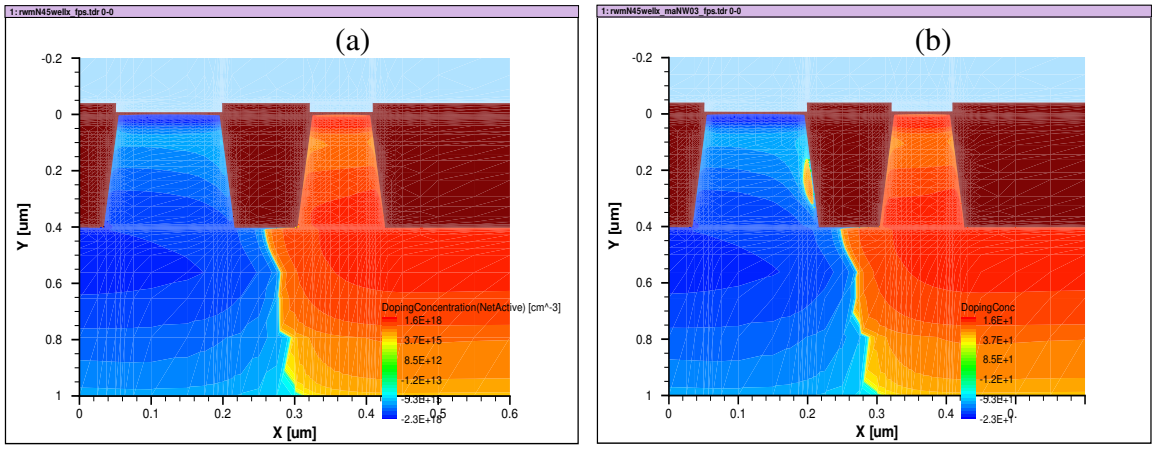
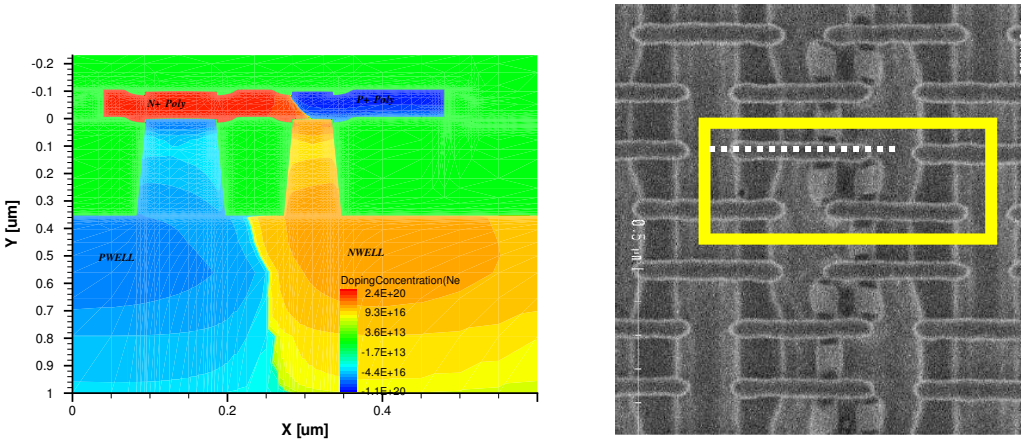


Figure 2-2 Simulated 45nm process with n+ to p+ spacing of 0.09µm. (a)Well doping profile contours of PWell (blue) and NWell (orange) with perfect NWell mask alignment and (b) with 30nm misalignment of the NWell.

The physical mechanism of lateral dopant straggle stemming from nuclear collisions of the high energy implanted ions in the photo-resist can have a relatively long range effect [30]. Because of this, it can impact both logic devices as well as the devices in the dense SRAM cell. Because of the well proximity for the bit cell devices, all the devices in the cell are subject to this mechanism regardless of the cell topology however only cell topology 4 would result in a within cell mismatch resulting from this mechanism. The implications of this proximity mechanism for the bit cell are two fold. First, this mechanism can produce a threshold voltage offset in the SRAM devices with respect to isolated logic devices and second, for the type 4 topology, it is an additional source of non-random mismatch. This amount of near surface doping is proportional to the dose of the high energy implant used in the formation of the retrograde well. As shown in Fig. 2-4, using implanted B11 energy of 200keV with a dose of 3E13 the near

surface doping is a function of the distance from the resist sidewall. Because the surface concentration is a function the distance from the resist sidewall, there is a clear potential for alignment sensitivity for the SRAM devices. Because of the relatively high channel doping in most nanoscale SRAM cells, provided deep retrograde implant doses are kept in this range, the impact of this mechanism on nanoscale CMOS SRAM will be limited.



(a) (b)

Figure 2-3 (a) Cross section TCAD simulation of type 4 cell topology illustrating the concern with poly inter-diffusion across the narrow n+/p+ space in the dense SRAM environment. Phosphorus doping has diffused into the channel region of the pull up PMOS device which will alter the work function and threshold voltage of the PMOS. (b) White dashed line indicates the cross section simulation region from SEM top view showing gate poly and active silicon regions of 45nm SRAM [31].

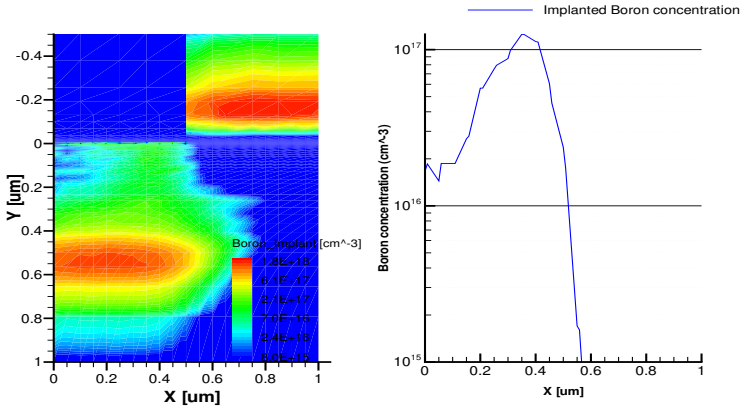


Figure 2-4 Atomistic Monte Carlo simulation of boron concentration across the silicon surface at a depth of approximately 50nm associated with lateral straggle. The resist is located at 0.5µm to 1µm on the X axis.

Photo-resist implant shadowing is another physical mechanism that becomes a factor in the dense SRAM design. The halo or pocket implant, to control short channel effects, is commonly implanted at angles in the range of 20-45 degrees. Because of the

pushed rules in the SRAM cell, the thickness and proximity of the photo-resist can result in implant shadowing in the dense SRAM devices. This can result in threshold voltage shifts in the SRAM devices relative to the logic devices and for the type 4 topology, can also introduce non-random mismatch.

2.3 Random versus systematic sources of variation

As discussed in chapter 1, Pelgrom's method of applying a Fourier analysis to separate the global variation sources from short range (mismatch) sources such as random dopant fluctuations (RDF) can be effectively applied to the SRAM devices. To perform this analysis requires a unique set of structures. The layout must be carefully controlled so that adjacent pairs of identically designed NMOS and PMOS transistors having sufficient variation in W and L values to enable a slope (A_{Vt}) to be extracted. These structures should be logic rule based, avoid proximity to resist edges and avoid potential lithographic related shape modifications due to effects such as corner rounding or foreshortening. By direct comparison of extracted A_{Vt} from the SRAM devices to that of the ideally drawn will provide a means of assessing the degree to which the systematic variation sources discussed in section 2.2 are present. Cell topologies which reduce the systematic sensitivities while maintaining the layout density and manufacturability advantages are clearly desired.

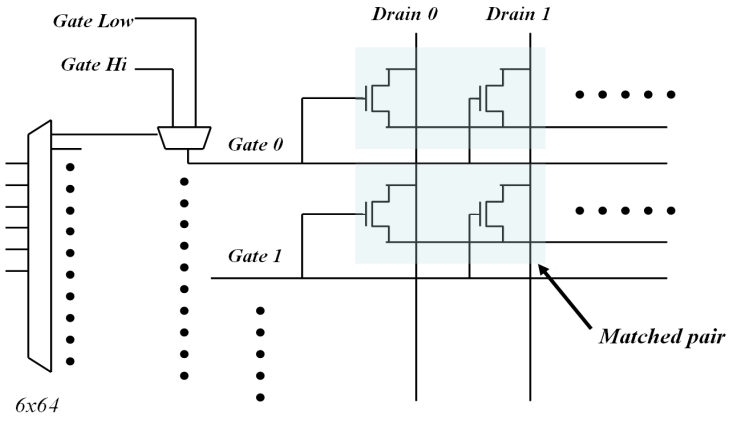


Figure 2-5 Schematic circuit diagram of device mismatch characterization circuit implementation to enable investigation of FDSOI 180nm devices.

A potential technology solution to provide improved A_{Vt} is FDSOI. To enable characterization of the benefits of this technology, a layout and circuit implementation of the mismatch structures has been completed in a 180nm FDSOI technology. A schematic of this design, which uses a 6x64 decoder to access 8 banks of NMOS and PMOS device pairs with the following (W/L) geometry matrix: W= 5, 1, 0.8, 0.6 μm and L= 0.15, 0.2, 0.6 and 1.5 μm is shown in Fig. 2-5. A full characterization and analysis of the hardware data is scheduled for 3/2010. One important point for further exploration is the recognition that the Pelgrom simple $(WL)^{1/2}$ model does not accurately capture the effects of randomness observed in today's CMOS technologies. Specifically, the threshold

mismatch values measured for larger L dimensions are observed to fall on a different slope than that of the minimum L [32] [33].

2.4 Cell design Compatibility with bias based assist methods.

As the functional margins are decreased with continued scaling, assist methods are becoming more essential for extending the functionality of the 6T SRAM cell. The complexity and feasibility for implementation of a specific assist method is dependent on many variables. The area overhead, cost design complexity, power and functional effectiveness as well as and cell topology are among the most significant factors. Because of the dependence on cell topology, this work addresses the impact of the cell topology in defining the optimum assist method.

The way the external supplies, WL and BLs are routed to and within the cell will strongly affect the types of assists method compatibility. The read assist options are first considered with cell topology 4. Cell topology 4 has a specific wiring design represented schematically in Fig. 2-6. Because the WL is on M3 and runs orthogonal to the BLs and the VDDc which are both on M2, the boosted VDDc read assist option would require that all columns along the asserted WL be boosted. Because the VSSc bus runs parallel with the WL, negative VSSc read assist could be employed with this cell layout, however, the functional sensitivity must also be taken into account as shown in Chapter 3. The read assist option of WL droop is a compatible option; however, performance penalties must be addressed.

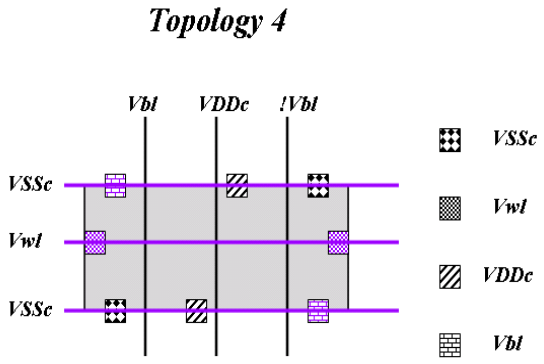


Figure 2-6 Wiring diagram used for type 4 topology cell.

2.5 Future Nanoscale Cell Topology

This proposed work explores the cell design options that are most promising beyond 32nm. Further exploration is proposed to investigate those 6T topologies that work most effectively with specific bias based assist methods. Finally, the role of technology based solutions such as the FINFET or FDSOI and non-6T topology alternatives will be evaluated.

Specific Contributions in the field of SRAM cell design to include:

- ***Identification of process technology interactions with SRAM cell design***
- ***Circuit design, hardware measurement and analysis of FDSOI device mismatch***
- ***Design and measurement of 6T, 8T and 10T array Iddq in FDSOI SRAM designs***
- ***Breakout of the design trade offs for competitive bit cell***
- ***Bit cell design criteria and optimization for incorporation of assist methods beyond 32nm***

Publications

i. ***SRAM cell design for stability methodology, VLSI Technology, 2005. (VLSI-TSA-Tech). 2005 IEEE VLSI-TSA International Symposium on 25-27 April 2005 Page(s):21 - 22***

ii. ***Noise Margin and Leakage in Ultra-Low Leakage SRAM Cell Design, IEEE Trans. Elec. Dev. Vol. 49, no. 8, Aug. 2002, p. 1499-1501***

iii. ***Optimization Criteria for SRAM Design- Lithography Contribution, Proceedings of SPIE - The International Society for Optical Engineering v 3679 n II 1999. p 847-859.***

Related Patents:

- *6,614,124 Simple 4T static ram cell for low power CMOS applications*
- *6,420,746 Three device DRAM cell with integrated capacitor and local interconnect*
- *6,967,351 Finfet SRAM cell using low mobility plane for cell stability and method for forming*
- *6,778,449 Method and design for measuring SRAM array leakage macro (ALM)*

CHAPTER 3

Coping with variability: SRAM Circuit Assist Methods

-Abstract

The 6T SRAM cell design has been successfully scaled in both bulk and SOI down to the 32/28nm node and has remained for more than a decade the dominant technology development vehicle for advanced CMOS technologies. Reduced device dimensions and operating voltages that accompany technology scaling have led to increased design challenges with each successive technology node. Large scale 6T SRAM beyond 65nm will increasingly rely on assist methods to overcome the functional limitations imposed by scaling and the inherent read stability/write margin trade off. An objective metric based methodology is developed for the evaluation of scaled CMOS technologies to provide guidance in the selection of assist methods. This paper explores various assist options given the technological constraints, functional boundary conditions and scaling trends that must be addressed for successful migration beyond 32nm.

3.1 Introduction

A unique feature of the 6T SRAM is an inherent balance between stability when holding data during a read or non-column selected write access and the ability of the cell to be written. This fact means that the device dimensions and threshold voltage targets established for the SRAM devices are a compromise by design. The ability to read and write will be characterized in terms of margins to assess the functional implications. These margins, will be referred to as write margin (WM), and read static noise margin (RSNM) or static noise margin (SNM), tend to decrease with scaling. When this fact is considered in context with the growth in bit count and increased variability with each successive generation, we may better comprehend the true nature of the mounting concern. This work seeks to explore the circuit options that may be needed to overcome the collapsing window of functionality and to provide a methodology for evaluation of the circuit assist options.

With continued scaling, circuit assist techniques will become necessary to preserve the 6T cell functional window of operation as scaling continues. A variety of SRAM functional assist methods have been proposed, however there remains no clear agreement in the industry as to which method or combination of methods will emerge as the more optimal solution. Moreover, different works compare the assist features in varied settings of technology node and technology type, but little detail is given on the trade offs involved in the selection process. Therefore, one goal of this paper is to develop an objective, metric based methodology to provide guidance for selecting an optimum assist feature for a technology platform. A second objective is to explore the impact of CMOS scaling trends on the robustness of various assist methods.

To address the first objective, a set of metrics are proposed and methodology is defined that can be used to evaluate a given assist technique or a set of assist techniques. The second objective is addressed by comparing how the metric values may be changing

with technology scaling. To accomplish this, the advantages and disadvantages of assist methods are quantified as a function of scaled SRAM across four technology generations from 65nm to 22nm. By addressing these objectives, an objective methodology is developed to examine the circuit assist methods which will enable the designer to most effectively compensate for these observed and predicted trends in 6T SRAM technology. The specific contributions of this chapter are:

- A concise overview of 6T SRAM assist options and a succinct categorization approach.
- An objective method to analyze and provide guidance in determining the specific assist approaches most suited for a given technology and memory application space.

3.2 Assist Categorization

As a foundation for a systematic approach in examining the range of circuit assist techniques a categorization approach is introduced of the various methods. Three distinct circuit level types or categories are proposed to address the reduced window of functionality for the 6T SRAM:

1. Altering noise source amplitude or duration through the access transistor,
2. Modification of the latch strength or voltage transfer characteristics of the latch inverters,
3. Avoidance or masking by design or architecture methods.

While category 3 is included for thoroughness and encompasses a range of approaches, including alternate bit cell selection (such as the 8T), ECC masking, or prohibiting the half select issue during a write operation [25], the scope of this initial work will focus on those methods of type 1 and 2. A summary of circuit assist methods is shown in Table I and can be categorized by the predominant assist type or category. The assist type listed in Table 3-1 provides a first order explanation of why the assist method is effective. While the category types are useful for quickly analyzing the various assist techniques, they are not necessarily exclusive of one another, and in several cases some degree of both are at work to influence the net assist effectiveness to be discussed in more detail in a later section of this chapter.

Table 3-1 Summary of SRAM Circuit Assist Methods with Predominant Assist Type

Read Assist	Type	Write Assist	Type	Terminal(s)
Raise VDD	2	Raise VDD	1	global [a]
Raise VDD at cell	2	Reduce VDD at cell	2	VDDc
Reduce VSS at cell	2	Raise VSS at cell	2	VSSc
WL droop	1	WL boost	1	WL
Reduce Q on BLs [b]	1	Increase (BL-BLB)	1	BL & or BLB
Weaken pass gate NMOS	1	Strengthen pass gate NMOS	1	array PWELL [c]
Strengthen pull-up PMOS	2	Weaken pull-up PMOS	2	array NWELL

[a] VDD applied to VDDc, WL, BLs and NW terminals

[b] reduced voltage or capacitance on BL

[c] Well bias also modulates pull-down NMOS device in most bulk technologies

The assist methods listed in Table 3-1 can be used in combination, and most can be implemented in either a static or dynamic mode. The categories can be further distinguished by the voltage terminal or terminals which are manipulated. For example a change in the WL voltage would involve modifying one voltage level while a change in the global VDD would involve changing the voltage on 5 of the 7 available terminals associated with the 6T SRAM cell (VDDc, NWELL, WL, BL and BLB). Increased global VDD is a unique type of assist method and will be discussed in more detail in a subsequent section. Modification of device threshold voltage by process change or by means beyond the control of the circuit designer, are outside the scope of this investigation.

Read assist methods refer to the set of circuit options that are used to either reduce the read noise source (type 1) or improve the cell stability (type 2) so that the cell remains stable during a read access. The methods which weaken read disturbance include reduced word line gate voltage [34][35][36][37][38], increased pass gate threshold voltage through body bias [39][40], and reduced bit line charge by lowering the voltage or capacitance [41][42][43][44]. The methods that are intended to improve the resilience of the latch are increased array VDD (VDDc) [35][45][46][47][48], decreased array VSS (VSSc) [36], and strengthening the pull up PMOS device by NWELL bias [39].

Those bias conditions which improve the write margin of the SRAM cell are referred to as write assist methods. These include boosted word line gate voltage [35][45][46][49] or reduced bit line voltage [36][38][50] to increase the V_{GS} of the pass gate device. These are referred to as type 1 write assist methods. Publications that address improving write margin by means of reducing the latch strength (type 2), include reducing the array supply voltage (VDDc) [34][35][37] [40][47][51], raising the array VSS (VSSc) [42][7][52] or reducing the strength the pull up PMOS by NWELL bias [39] [53].

3.3 Assist Functional Sensitivity

The metrics identified for assist evaluation are relatively simple: effectiveness, performance, power, and cost. The incorporation and commercial success of a given assist technique ultimately depend on these four fundamental metrics, although the relative importance may vary with the specific application. In this section the metrics which are essential to the proposed method are developed for evaluation of assist techniques.

Functional effectiveness is defined as the margin sensitivity to the change in voltage for a given technique. This is expressed as:

$$Sensitivity = \frac{\partial(Margin)}{\partial V} \tag{3–1}$$

Margin may refer in this case to either SNM or WM. To compare the functional effectiveness or margin sensitivity of the specific assist methods, noise margin analysis is performed using custom predictive technology models using pre-defined scaled SRAM dimensions consistent with the dense SRAM published values as shown in Chapter 1. As part of the methodology defined in this investigation, particular emphasis is placed on the specific conditions that represent the worst case operation voltage (Vwc) for the technology. Vwc is defined as the minimum voltage at which the SRAM must be able to

perform both a read and write operation across the entire array without failure. Thus, one must ensure that the VDDmin for a given array is at or below our predefined Vwc for each technology node. Because Vwc is recognized as technology and application dependent, 0.8X the nominal VDD will be used as this value. This condition accounts for factors such as voltage droop, NBTI shifts over the product lifetime, and testing equipment variability.

3.4 Assist Characterization

Four read assist and four write assist methods to provide a set of test cases for the assist evaluation methodology. Simulations are used to evaluate the impact of specific assist methods and characterize the impacts on stability and write margin to assess the functional sensitivity of the methods. The performance was evaluated following the previously defined methods using the simulations of read current and the cell write time. Finally, the impact of the assist method on variation was evaluated in order to fairly compare the true efficacy of an assist technique in improving the yield for a large scale SRAM array.

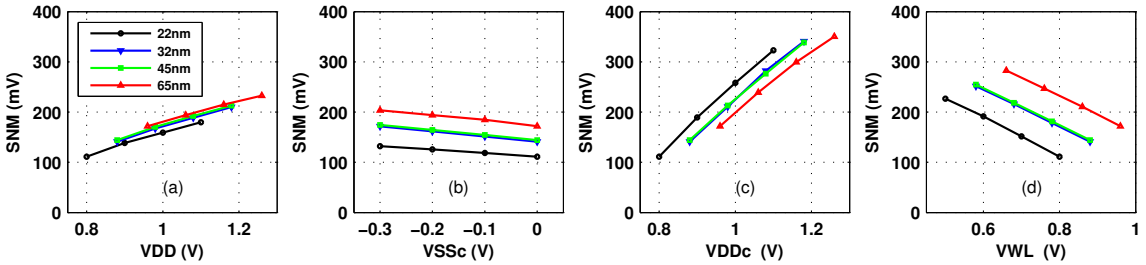


Figure 3-1 Static noise margin as function of (a) raised array global VDD, (b) Negative VSS at the cell, (c) VDD boost at the cell and (d) WL droop.

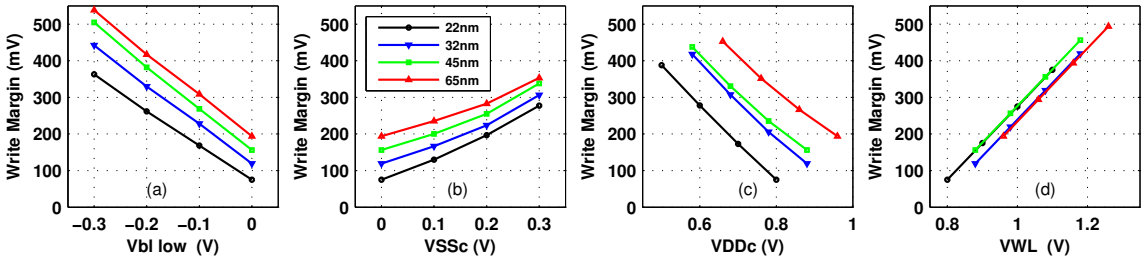
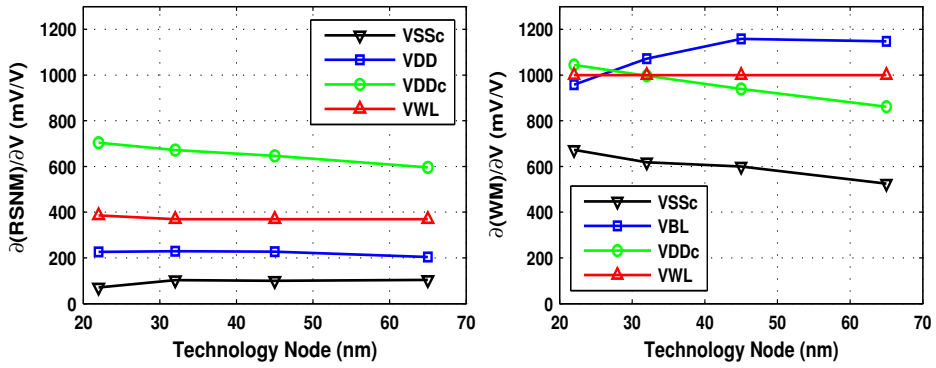


Figure 3-2 Write margin as function of (a) negative BL, (b) raised VSS at the cell, (c) VDD droop at the cell and (d) WL boost.

Fig. 3-1 plots the SNM as a function of the assist bias for the four read assist techniques evaluated. The four technology nodes are represented in each of the four plots. Fig. 3-1(c) for example shows the change in SNM with increased array VDD (VDDc). There is a negative slope for methods (b) and (d) corresponding with the fact that these methods require a reduction in the terminal voltage. While all four methods produced some degree of improvement in the SNM, and the response or sensitivity is similar across the technology nodes, the sensitivity was most significant for VDDc. The initial voltage is either 0V or varies consistently with the Vwc for each technology.

The simulation results for the write assist methods are shown in Fig. 3-2 (a-d). For the write assist methods in this analysis, the VSSc response, Fig. 3-2b, was the least

linear and showed the least sensitivity. Although there is some degree of non-linearity in the response characteristics of write margin and static noise margin, most exhibit a sufficient degree of linearity across the 300mV range to allow us to characterize the responses using a first order linear model to allow a high level comparison. SNM sensitivities shown in Fig. 3-1 (a-d) are summarized for each of the technology nodes in Fig. 3-3(a). As a means of improving the SNM, raised cell voltage (VDDc) is the method that emerges as exhibiting the greatest sensitivity across the LP technology nodes. The trends also suggest that there is some increase in sensitivity as scaling continues.



(a) (b)

Figure 3-3 The margin sensitivities across LP technologies for the four read assist methods (a) and four write assist methods (b) investigated.

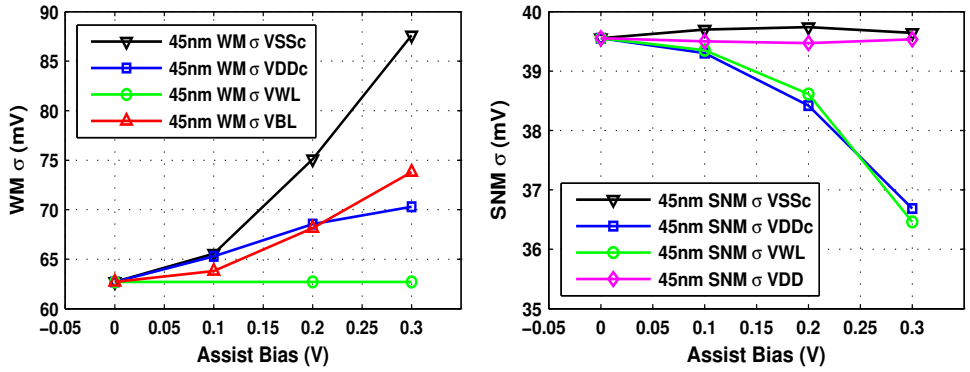
Following a similar approach, the functional sensitivities were also characterized across the technology nodes for write margin sensitivity, Fig. 3-3(b). In this case, three of the methods exhibit similar sensitivities to the applied bias. Raised array VSS (VSSc) showed less degree of linearity and had a weaker response. The unique and completely linear response of the WL boost was due to the fact that the write margin metric used in this investigation was defined as the difference between the final word line voltage and the voltage of the word line required to write the cell.

3.5 Yield

Until now, only the impact of SNM and WM mean values at a given bias condition have been discussed for the assist methods. However, to determine the functional yield expectation for a given array size at the worst case voltage, the local and global variation must be taken into account. Without the variation component, the required margin improvement will be unknown. For the small scaled SRAM devices, the local variation associated with random dopant fluctuations (RDF) dominates the variation sources. Although technology improvements offered by high-κ and metal gate, may provide significant improvement due to the higher gate capacitance, continued scaling will quickly consume these gains.

The impact of the assist method on the variation in SNM and WM is now addressed. To investigate this, Monte Carlo simulations were performed for the four write assist and read assist methods for each technology node explored. Fig. 3-5 plots the sigma for the WM distribution (a) and SNM (b) as a function of the assist voltage bias for the 45nm node. A minimum of 200 Monte Carlo runs were performed for each bias

condition. Several observations emerge from this analysis. First, the assist method and bias impacts both the standard deviation of the distribution. This is accounted for in assessing the overall contribution of the assist method which will be discussed in the next section.



(a) (b)

Figure 3-5 Impact of assist method applied bias on the sigma of the resulting 45nm LP technology distribution for write assist (a) and read assist (b).

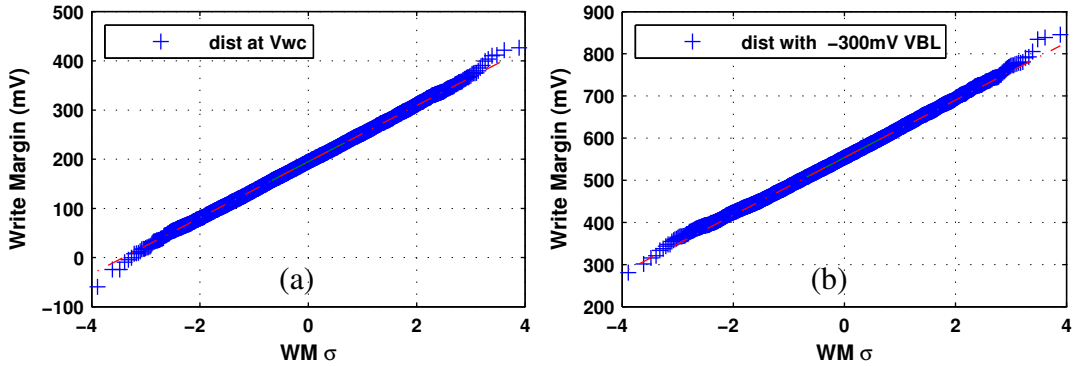


Figure 3-6 10,000 Monte Carlo cases showing WM(0) standard normal distribution for 45nm LP technology at Vwc with no assist bias (a) and with 300mV negative BL bias (b).

SNM0/WM0 is denoted as the read/write margin for data '0'and SNM1/WM1 as the margin for data '1'. The true definition of SNM/WM would be the minimum value for '0' and '1'. An important observation is that the distribution of SNM0 or SNM1 can be well fitted to normal distribution under normally distributed parameter variation. For the cases examined, the distributions remain normally distributed with assist bias, though the mean μ and the standard deviation σ may change. An additional set of Monte Carlo simulations (1,000 to 10,000 cases) were run on selected assist bias conditions for distribution verification purposes. Fig. 3-6 plots the results of 10,000 cases for WM0 at Vwc and with 300mV negative BL bias for the 45nm LP technology. The linearity of the quantile plots confirms that the WM distribution remains normal even with the assist feature engaged.

Because of the stochastic nature of the RDF induced variation and because the focus will be on large arrays with relatively few fails, the Poisson distribution will be

used as an approximation of the binomial distribution. The failure probability for the right or left node (probability of SNM0<0 or SNM1<0 for example) is given as:

$$Pf = \frac{1}{2} \cdot erfc\left(\frac{\sigma}{\sqrt{2}}\right) \tag{3-2}$$

The array fail probability (λ) is computed by including both states of the latch:

$$\lambda = N \cdot \left(Pf_{(0)} + Pf_{(1)}\right) \tag{3-3}$$

The array fail probability can then be expressed as a function of the distribution sigma (σ) of SNM0 or SNM1 for example and the number of bits in the array (N):

$$\lambda = N \cdot erfc\left(\frac{\sigma}{\sqrt{2}}\right) \tag{3-4}$$

With the assumption that the RDF induced variations are random and non-clustered, the soft fail yield for a given mechanism becomes:

$$Yield = \exp(-\lambda) \tag{3-5}$$

Following this approach, to obtain a 10M-b SRAM with a SNM-limited yield of 99% would require a sigma value of ~6.12σ. SNM0wc represents the worst SNM tail in the distribution that satisfies $Pf_{(0)}$ = 4.68E-10. To achieve this yield target, SNM0wc must be larger than the minimum noise margin threshold (e.g. 0) for 99 of 100 10M-b arrays. The limited yield for WM is computed with this same approach, to obtain a 99% WM-limited yield, which would result in an over all soft fail limited yield of 98% considering both WM and SNM. For the 45nm LP technology, Fig. 3-7 shows that this is achieved with 100mV assist bias for the most effective read assist technique (VDDc boost) and 180 mV for either word line boost or negative BL bias, which showed very similar sensitivity for the 45nm node.

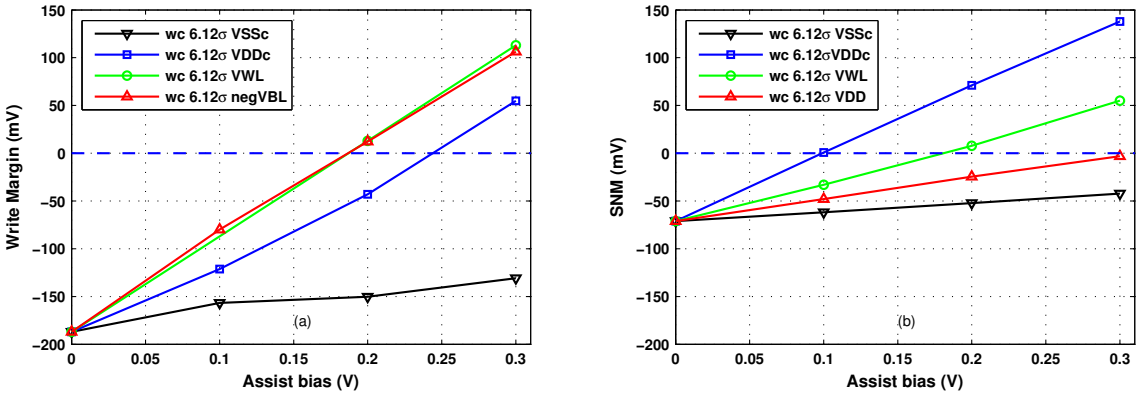


Figure 3-7 The 6.12sigma worst case write margin (a) and SNM (b) as a function of assist bias for the 45nm LP technology.

3.6 Discussion

The elements of a metric based assist selection methodology which is referenced to Vwc have been outlined. Additionally, a means of transforming the write and read margins into a yield assessment is provided. This approach has been applied and demonstrated using the LP PTM platform of bulk technologies from 65nm to 22nm. The metrics considered fundamental are: margin sensitivity, performance, power and cost.

While specific circuit level implementation schemes are not provided, the focus of this effort was to provide a clear method for categorizing, evaluating and comparing the assist method(s) that should be incorporated for a given technology or technology platform.

The functional read/write margin sensitivity was evaluated over a 300mV window to minimize non-linearity in the response and to ensure the bias conditions would not exceed the technology reliability limits. Because the reference (Vwc) condition was more than 200mV below nominal VDD in all cases, the reliability requirement was preserved. Even for the 22nm node where the Vwc was taken to be 0.8V, the max voltage would be only 10% greater than nominal VDD, which is consistent with common technology specifications.

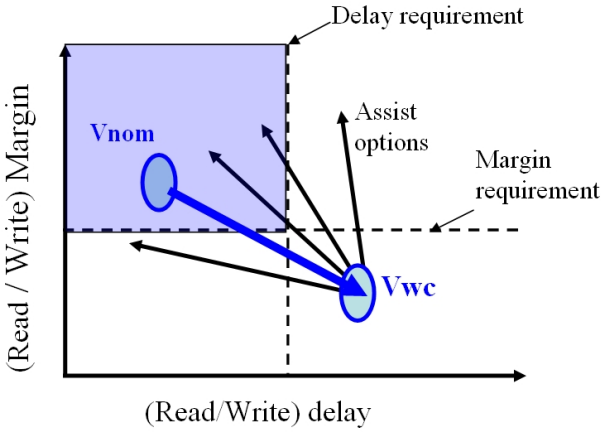


Figure 3-8 Schematic diagram of read/write margin vs read/write delay and desired functional window based on margin limited yield and performance requirements for application.

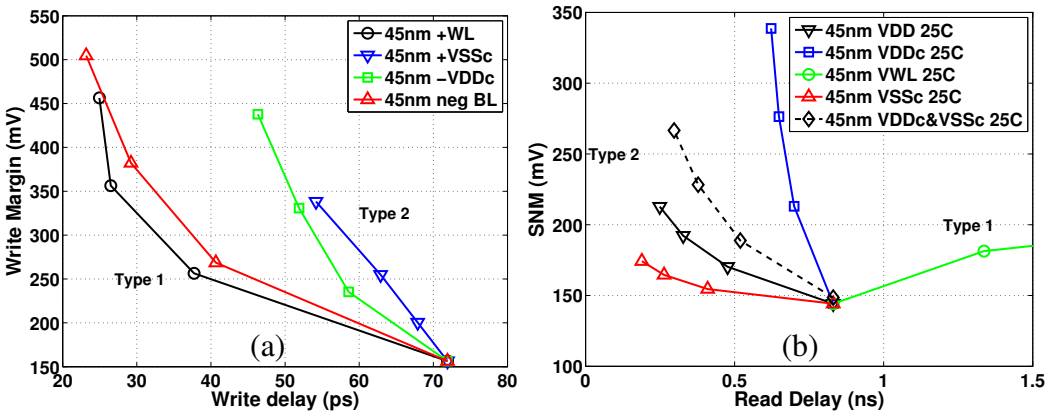


Figure 3-9 Margin vs delay plots showing write (a) and read (b) for the 45nm LP technology when assist bias is swept from 0 to 300mV.

The sensitivity response for the assist methods studied is often influenced by more than one mechanism and can be understood when the device physics are taken into account. For example, the superior result associated with raised array voltage (VDDc) for read assist can be attributed to the fact that several mechanisms influence the result. The body effect causes the cell PFET to become stronger because of the modulated VSB

for the PFET and the VGS is increased for the devices in the latch which are being held on.

A method to estimate the performance for both the read and write operation was discussed, and the essential components of performance were identified. The normalized margin sensitivity and performance as provided in Table II is summarized to obtain an overall functional effectiveness (final) factor. The product of the normalized values, assuming equal weighting for each (different weight factors could easily be applied if desired), provides a simple but effective technique to rate the effectiveness of an assist feature. Considering the combined effect of both margin sensitivity and performance impact leads to an interesting observation regarding the merits of the two assist categories. For the techniques evaluated in this paper, the type 1 methods emerge as the more advantageous for write assist while the type 2 methods tend to be preferred for read assist.

For those applications where both the margin limited yield and performance requirements are known, an alternate (margin/delay) method of establishing the viable assist method(s) and corresponding bias conditions is proposed. A schematic diagram depicting the desired functional window, delineated by the margin and delay requirements of the memo is shown schematically in Fig. 3-8. As the VDD is reduced to Vwc, the read/write margins and corresponding performance degrade. Use of assist methods generally improve margin and in most cases delay to some extent. The margin/delay analysis illuminates the most effective assist methods for a given technology and functional requirements and provides additional insight into the net functional impact of a given assist approach. This graphical approach allows the designer to quickly realize the essential functional advantages of a given assist method or combination of methods. An example of the margin/delay plot for the 45nm LP technology is shown in Fig. 3-9 (a) for write margin and (b) for read margin. A combined VDDc and VSSc assist method is shown in Fig. 3-9 (b) which demonstrates how the assist techniques can be combined as required to optimize both delay and margin. This figure also points out that some assist methods such as WL droop may improve the margin while simultaneously degrading the performance.

The effect of variation was examined in some detail and it was found that both assist method and bias had a non-negligible impact on the resulting WM and SNM distributions. For those cases where the assist method influenced the distribution, it was necessary to account for this in determining the effectiveness of a given method on the yield. While the SNM and WM distributions are intrinsically non-Gaussian for reasons previously discussed, an approximation of the true distribution is used to compute the distribution tail. By this method, a required assist bias for a given array size and soft fail yield requirement may be established for both WM and SNM.

3.7 Evaluating Power and Cost

Upon establishing the functional effectiveness of the assist methods, the power and cost implications for the methods which meet the required functional objectives may be considered. The impact on power associated with an assist method will largely depend upon the specific implementation scheme. For example, most of the methods discussed can be implemented in a dynamic as well as static form, and these types of implementation choices will have a large impact on power. Because the effectiveness can

be considered a necessary condition for incorporation, the initial efforts are focused on quantifying the sensitivity and performance impacts of the method. The impacts of cost and power are then considered only for those methods that may provide sufficient functional margin.

Cost remains a challenging metric to fully quantify, and several cost factors are discussed that may be considered in selecting an assist method: area, complexity, scale-ability and compatibility. The area for implementation of a specific assist scheme is typically 2-4%, while those methods that involve well bias are considerably higher. For complexity, any yield implications associated with implementation are included into this cost category, but it could also be considered separately. For example, of the four write assist methods investigated, three (WL or VSSc boost, and VDDc droop) require a higher, yield related, complexity of implementation. This is because WL boost raises the potential of stability upset in the cells on the non-selected columns, and reduced voltage at the cell by VSSc boost or VDDc droop introduces data retention concerns. Although the negative BL method avoids this yield side effect, the cost of generating the negative voltage and level shifters must be taken into account.

Table 3-2 Summary of Sensitivity and Performance Factors for the Subset of Assist Methods Examined

	Method	primary Type	Sens factor	Perf factor	final factor[a]
Read Assist	Raise VDD	2	0.329	0.799	0.26
	Raise VDDc	2	1.0	0.160	0.16
	Reduce VSSc	2	0.100	1.0	0.10
	WL droop	1	0.557	0.002	0.00
Write Assist	Reduce VDDc	2	1.0	0.24	0.24
	Raise VSSc	2	0.606	0.23	0.14
	WL boost	1	0.962	0.947	0.91
	negative BL	1	0.923	1.0	0.92

[a]based on equal weighting of margin sensitivity and performance

The scale-ability cost factor was addressed by examining the low power technology platform with custom PTMs. For this case, there was some degree of consistency across the technology nodes and those methods which were most effective at the 45 and 65nm nodes remained so through the 32 and 22nm nodes. For the final cost category, compatibility, both cell design and assist method compatibility should be considered. For assist method compatibility, there are advantages when certain assist methods can share a common supply. For those applications in which power is less critical than cost, increased static supply (VDD) by separate voltage islands may be the most cost effective method. This approach can allow the logic to operate at a lower voltage than the memory.

For applications where power is more critical, the dynamic implementations are more favorable with a cost of some degree of added complexity. Regarding the cell compatibility with a given assist method, it is noted that the 6T cell is typically provided

by the foundry and therefore constrains the memory array designer to seek assist methods that best comply with the given layout. For example, the predominant industry 6T cell design style makes use of a VDD bus on metal level 2 (M2) level running parallel with the M2 bit lines. Although this layout style has advantages for density and performance reasons, the implementation of locally raising VDDc along the word line requires that all columns on the selected WL be boosted. Although pulsing the VSSc may be more consistent with this style cell layout on the metal 3 (M3) VSSc bus which runs parallel with the M3 word local line, this technique exhibits less margin sensitivity. It should also be pointed out that assist compatibility with dual port SRAM is of emerging importance, and some methods such as drooped VDDc for write assist are fundamentally incompatible. For those applications requiring both 1 and 2 port SRAM, the cost effectiveness for an approach such as the negative BL may become more compelling.

3.8 Analysis

Following this proposed method, the sensitivity and performance factors are summarized for the set of assist methods examined as shown in Table 3-2 or if specific requirements are known following the approach shown in Fig. 15. For those methods evaluated with the PTM technologies the selection process is continued by summarizing the cost impact for those methods most effective based on margin sensitivity and performance. Along with raised global VDD, four additional combinations of assist methods would need to be considered. If one chooses to consider those assist method combinations for the predominant industry cell layout, the cost comparison may then be summarized in Table 3-3. For the LP bulk technologies considered in this study, both read and write assist would be required to achieve high yield for large SRAM arrays beyond 65nm. Based on the margin sensitivity and performance metrics (equally weighted in this investigation), five potentially viable combinations of assist options were identified. By introducing the additional constraint that the 6T cell must be consistent with the bit cell layout most commonly used in the industry, the viable assist combinations reduce to three. For a final selection between the combinations of assist methods, the power metric is considered. For power sensitive memory applications, the power consumption required by the increased global VDD across the array would be less attractive. However, power reduction methods such as aggressive use of sleep and or drowsy mode may sufficiently mitigate this penalty while introducing some additional complexity.

For the given subset of assist methods examined using this metric based method, those which appear most viable for future development in LP bulk CMOS technologies were successfully identified. For cost sensitive LP bulk technologies the assist methods that will receive more focus will be memory voltage islands with raised VDD, and negative BL combined with either negative VSSc or VDDc boost for write and read assist respectively (Table 3-3 rows 1, 3 and 5).

Based on the observation that the negative VSSc has a weaker sensitivity for improving RSNM but exhibits improved read performance, a new read assist approach is proposed for the type 4 topology (discussed in chapter 2), invoking a combination of WL droop and negative VSSc. Similar reasoning is used to address the options for write assist. Drooping the VDDc may induce data retention failure in the cells along the drooped column that are not being written. VSSc boost, asserted in parallel with the WL

can be used but a yield impact exists for the half-selected cells boosted VSSc. WL boost will likewise share the same concern with the half-selected cells. Using the negative BL with the type 4 topology is therefore the most compatible option without invoking array architecture modifications.

Table 3-3 Cost Factors for Viable Read and Write Assist Combinations

Read Assist Method	Write Assist Method	cell compatible	low yield complexity
Raise VDD	Raise VDD	yes	yes
-VSSc	+WL	yes	no
-VSSc	-BL	yes	yes
+VDDc	+WL	yes[a]	no
+VDDc	-BL	yes[a]	yes

[a]VDDc boost required for all columns on asserted WL

Following this analysis, a potential synergy exists in the development of the negative voltage can be used for both read and write assist. It is therefore concluded that a combined set of assist methods for read (negative VSSc with drooped WL) and negative BL for write provide the most effect functional benefit for future, type 4 topology, 6T designs while preserving the conventional array architecture.

3.9 Conclusion

As competitive forces and industry scaling continue to erode the 6T SRAM functional margins, the use of assist methods will increase. A review and categorization approach for examining potential assist methods is provided. Based on the set of assist methods evaluated, for the LP bulk CMOS technologies those methods classified as predominantly type 1 more effective for write assist and the predominantly type 2 category of assist methods more effective for read assist. A methodology was developed and demonstrated to provide an objective and comprehensive means of establishing the optimum 6T SRAM assist method for a given technology and set of application requirements. The objective, metric based methodology proposed defines an approach which: (1) is referenced to Vwc, (2) defines and quantifies the critical metrics, and (3) provides a means of transforming the write and read assist margins into a yield assessment. As shown using LP PTMs, this metric based approach can be successfully applied to provide guidance in selecting the optimum set of assist methods for a given technology platform. Based on the analysis presented in this paper, future bit cell design and development efforts that incorporate preferred circuit assist methods, will be significantly advantaged.

Specific Contributions in the field of SRAM circuit assist methods to include:

- ***Methodology to provide guidance for the circuit designer in the selection of assist methods for SRAM that includes:***
  - ***Assist categorization approach***
  - ***Analysis and margin sensitivity centered at the worst case voltage***
  - ***New margin/delay space analysis for evaluating assist methods***
- ***Quantification of SRAM performance and functional margin across LP CMOS technology platform (65nm-22nm)***

Publication @UVA:

i. Methodology for Evaluation of 6T SRAM Assist Methods in Scaled CMOS Technologies (to be submitted for publication)

Related Patents:

- *7,313,032 SRAM voltage control for improved operational margins*
- *7,075,153 Grounded body SOI SRAM cell*

CHAPTER 4

Limits of Bias Based Circuit Assist Methods in Nanoscale SRAM

-Abstract

Reduced device dimensions and operating voltages that accompany technology scaling have led to increased design challenges with each successive technology node. Large scale 6T SRAM arrays beyond 65nm will increasingly rely on assist methods to overcome the functional limitations imposed by increased variation, reduced overdrive and the inherent read stability/write margin trade off. Factors such as reliability, leakage and data retention establish the boundary conditions for the maximum voltage bias permitted for a given circuit assist approach. These constraints set an upper limit on the potential yield improvement that can be obtained for a given assist method and limit the minimum operation voltage (Vmin). By application of this set of constraints, it is shown that the read assist limit contour (ALC) in the margin/delay space can provide insight into the ultimate limits for the nano-scale CMOS 6T SRAM.

4.1. Introduction

Increased device variability and reduced overdrive associated with lower operating voltages have reduced the functional yield margins in VLSI circuits. This is particularly true for the 6T SRAM, which continues to play a dominant role in future technology generations because of its combination of density, performance, and compatibility with logic processing. The successful commercial scaling of the 6T SRAM driven by strong industry competition is expected to continue beyond the 32nm node. The continued trend in area reduction is accompanied by the well known consequence of increased variability associated with the reduced channel area. Although technology options such as high-κ with metal gate may provide some relief in variability, diminished functional margins coupled with the growth in bit count pose a serious technical challenge beyond the 28/32nm generation.

Because of the commercial success of the 6T SRAM, methods to address the failure mechanisms of large memory arrays will extend the life of the 6T SRAM in VLSI circuits. Fail types for SRAM arrays may be divided into two distinct categories: “hard fails”, i.e., those attributable to defects, and “soft fails”. Soft fails defined in this context are those voltage, temperature and timing dependent fails resulting from one of the following four modes: (1) failure to write, (2) failure to read (insufficient signal developed on the BL), (3) stability upset during a read or half-select condition, and (4) data retention failure. These four failure modes each first occur at the distribution tail stemming from global and local variation sources.

The use of bias based circuit assist methods has become increasingly common, primarily to address soft fail modes 1 and 3 and to preserve the 6T cell functionality as the variation continues to increase and both read and write margins decrease with scaling. Although numerous recent articles have discussed bias based assist for SRAM as reviewed in chapter 3, limitations exist for all of these techniques. This limit may be

reliability, performance, leakage, energy, or other factors which ultimately bound the extent to which the assist method compensates for the reduced functional margins.

The objective of this chapter is to explore the boundaries of bias based assist methods to understand the impact on the minimum operation voltage (Vmin) and the effectiveness of the assist methods for future generations of 6T SRAM. Based on the relationship between performance and functional margin with the applied bias constraints, the assist limit contour (ALC) for read assist is defined across four technology generations. For write assist methods, besides the constraint from reliability, the read stability of half-selected cells limits the permissible assist bias.

By application of the constraint limitations the maximum assist margin values permissible can then be mapped. The maximum permissible assist bias based on the reliability constraints are defined for each technology. The reliability limit may be due to several factors such as time dependent dielectric breakdown, hot carrier, NBTI or a combination of the known mechanisms with sufficient voltage acceleration. The maximum assist bias offset $|V_{assist}|$ that may be applied for any given assist method based on the reliability (Vmax) constraint is expressed as:

$$\left|V_{assist}\right| = (V_{\max} - V_{nom}) + V_{droop} \tag{4-1}$$

V_{nom} and V_{max} refer to the nominal and maximum operation voltage as specified by the technology developers (Vnom values provided in Table 1-1 of chapter 1 for this work). V_{droop} refers to the difference between V_{nom} and the instantaneous operation voltage. To illustrate this concept briefly, for a technology in which the Vnom/Vmax is 1.2V/1.32V respectively, if the array VDD is drooped from 1.2V to 1V, the maximum assist bias is 0.32V. Any bias exceeding 0.32V would exceed Vmax for the transistor, violating the reliability constraint. For the same reason, a maximum negative assist bias of 0.32V may be applied provided all VDD supply terminals associated with the array are maintained at 1V. Additional constraints may apply, but this single constraint provides a defined boundary that will be discussed further in section 4 of this chapter.

Table 4-1: Summary of constraints for bias based assists

Assist	Method	Bias Constraint
Read Assist	WL voltage ↓	Vmax
	Pass Gate Vt ↑	Vmax
	BL voltage ↓	Vmax and Write 0
	Array VDD ↑	Vmax
	negative VSS ↓	Vmax,Vfwd
	PMOS |Vt| ↓	Vmax
Write Assist	WL voltage ↑	Vmax, RSNM (½ Select)
	negative BL ↓	Vmax and Vfwd
	array VDD ↓	Vmax, DR (with shared VDDc)
	Array VSS ↑	Vmax, DR (with shared VSSc)
	PMOS |Vt| ↑	Vmax, RSNM (½ Select)

In addition to the technology defined Vmax constraint, other assist bias constraints for read assist bias include; forward bias diode turn-on (Vfwd) when VSSc is intentionally driven below ground, and cell upset by writing a zero when both bit lines are drooped sufficiently low (Write 0). For write assist, the constraints are again reliability (Vmax) as well as data retention (DR) for non-accessed cells sharing the intentionally modulated common supply, forward biased diode turn-on (Vfwd) when the

'write zero' bit line is driven below ground, and cell stability for the half-selected cells on the asserted word line. The primary bias constraints are summarized in Table 4-1 for the bias based assist methods evaluated in this paper. Vmax is a valid constraint for all cases. This is less obvious for the two write assist options that involve collapsed supply across the latch. Vmax remains a constraint for the maximum write margin because it still limits the maximum WL voltage.

For the purposes of this work, Vmax will be defined as 10% above the nominal operation voltage. Because Vmax is a limiting factor in all bias based assist methods, this fact may be exploited to explore the limits of the assist methods across the scaled technologies. This approach allows us to effectively define the upper envelope of assist bias conditions permissible for a given technology. By mapping the assist methods across the margin/delay space, the functional window may then be used to illuminate the practical voltage bias boundaries.

4.2 Results

To examine the maximum soft fail limited yield boundaries that can be achieved for a given assist method, the relationship with VDD is first described and then applied to the assist bias using the Vmax constraint. The read static noise margin as a function of VDD is shown in Figure 4-1(a) for the 45nm LP PTM technology. The three cases shown are with array VDD (VDDc) boost, array VSS (VSSc) reduced, and with no assist. It is clear from Figure 4-1(a) that the use of the maximum assist bias, consistent with relationship (4-1), can significantly improve the otherwise reduced static noise margin when the word line is asserted. The SNM improves beyond the nominal value when the VDDc assist method is invoked because the noise source is being reduced with VDD reduction, and the latch strength is increasing with the boosted VDDc.

Competing mechanisms produce a different result with negative VSSc. In this case, although the net latch strength is improved over the non-assist case, the noise source through the pass gate NFET is becoming stronger due to the body effect producing a reduction in pass gate V_T on the side of the cell storing a zero. Additionally, the V_T is reduced for the pull down NMOS device with drain storing a one. This results in an earlier turn of this pull down NMOS and further reduces the SNM. The read delay for the cell is improved due the body effect which strengthens both the pull down and pass gate series devices on the side of the latch storing a zero.

While the cell stability compensation is larger for VDDc assist, the improvement in performance (read delay) may not be sufficient depending on the functional window as discussed earlier. Boosting the VDD at the cell (VDDc) has a small impact on the read delay, consequently the read delay continues to degrade as VDD is reduced. The alternate read assist method (VSSc) shown in Figure 4-1 improves SNM to some degree but more significantly improves the read performance. This is because the body effect associated with reduced VSSc causes both the pull down and pass gate NFET device V_T to be reduced, boosting the read current.

The margin/delay relationship is applied for the assist methods with maximum assist bias. The effect of maximum assist bias on both SNM and delay based on the modulation of single and multiple terminals is shown in Figure 4-2(a) for the 45nm LP-PTM technology. Each of the read assist bias conditions given in Table 4-1 except those involving well bias V_T modulation were employed.

The margin/delay analysis reveals the limits of the bias based assist methods across the relevant design space. This boundary further defines a contour, as shown by the solid continuous line (demonstrated using VDDc and VSSc assist bias following the Vmax constraint). The boundary referred to as the assist limit contour (ALC). It establishes the effective limit in SNM and corresponding relationship to read performance for a given technology and bit cell. This boundary or ALC mapped by the assist methods therefore provides a means of assessing the functional limits of the 6T SRAM.

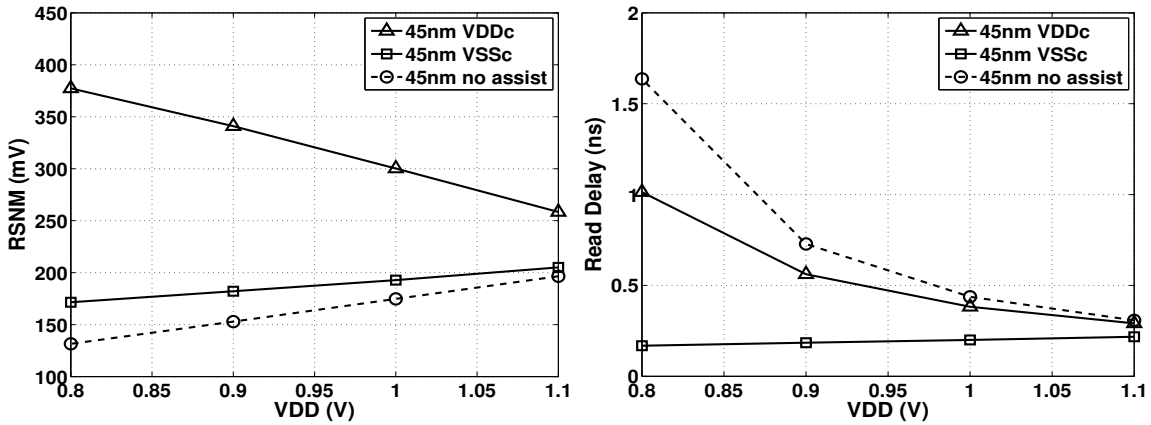


Figure 4-1 Change in RSNM with reduced VDD (a) and effect of VDD on read delay (b) with the maximum allowable assist bias at each VDD. Data based on 45nm LP PTM.

For the cases studied, the read ALC as defined by the latch supply voltages were found to provide a reasonable approximation of the full multi-terminal Vmax read assist contour. Because drooping the WL provides a degree of freedom that is not limited by the Vmax constraint, those combinations of negative VSSc combined with WL voltage reduction where found to produce a slightly improved margin/delay response for the LP-PTM technologies.

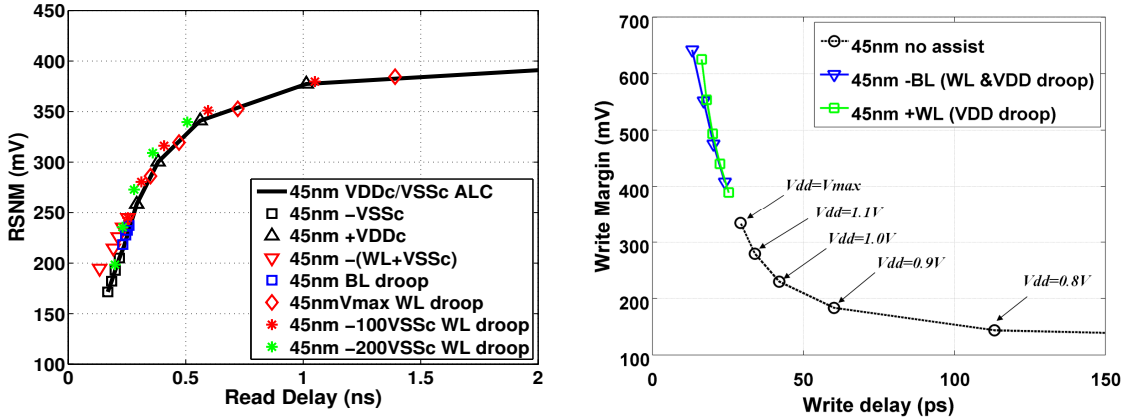


Figure 4-2 (a)Multiple read assist options involving both single and multiple terminals with Vmax constraint preserved. (b) Write margin decreases as VDD is reduced

when no assist is used. With assist at Vmax, the write margin is increased with reduced VDD.

Because the primary goal of this work is to identify and delineate the bias based assist limitations of the scaled 6T SRAM cell, the delay required in developing the bias conditions is not included in this analysis. A complete SRAM macro design would need to include the overhead delay associated with the specific implementation and circuit choice.

For write assist, the margin/delay analysis leads to a different result. In this case, there is no inherent trade off between write margin and write delay. The relationship is shown in Figure 4-2(b) for two assist methods (negative BL and boosted WL) showing improved margin and delay as array VDD is drooped. For boosted WL assist, the array VDD, high bit line and NWELL voltages are reduced, while the WL line is boosted to Vmax limited by the reliability constraint between the WL voltage and the low (write zero) bit line at 0. For the negative BL case, as the VDD is reduced on the word line, high bit line, and array VDD, while the low bit line is drooped by the same amount to preserve the Vmax constraint.

With the word line boosted to Vmax, the write margin continues to increase with corresponding VDD reduction. Similarly, with the (write zero) bit line driven below ground by a value equivalent to the VDD reduction (preserving the Vmax constraint), the write margin continues to increase. In addition to Vmax, the maximum write assist bias may become limited by other constraints, such as the read margin for the half-selected bits, shown in Table 4-1.

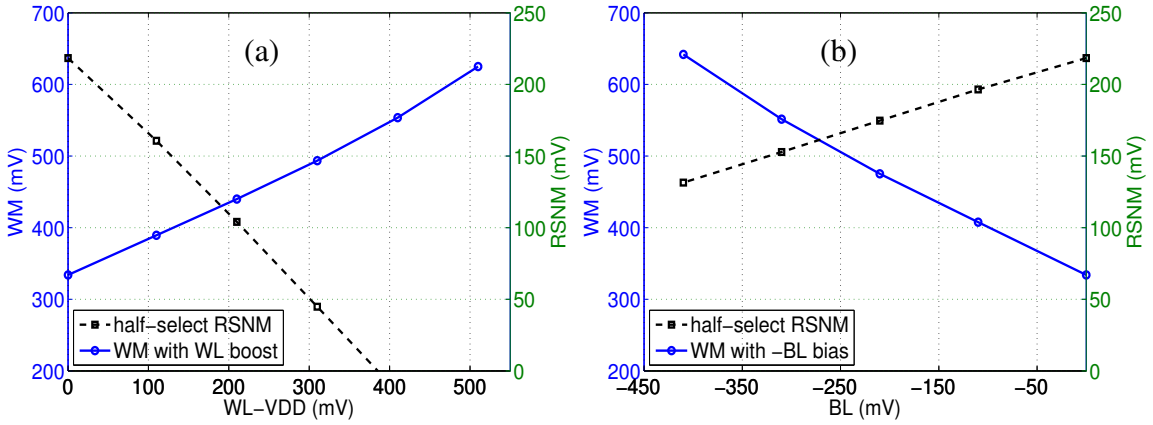


Figure 4-3 (a) The impact of WL boost on the WM of the selected bits and the stability (RSNM) of the half-selected bits. (b) The impact of negative BL on the WM of the selected bits and the stability (RSNM) of the half-selected bits.

For the WL-boost write assist, Figure 4-3(a) shows that as the array supply voltage is reduced, boosting the WL while preserving Vmax reduces the stability (RSNM) of the half-select bits on the same WL. Therefore, the limiter for the WL boost quickly becomes the reduced SNM on the half-selected bits. For the negative BL assist, the BL bias does not directly impact the half-select bits. However, because the amount of bias between the

BL voltage and the global VDD is limited to Vmax, a larger negative bias on the BL implies a lower global VDD. Thus the RSNM of the half-selected bits consequently decreases, permitting a larger negative BL bias, as shown in Figure 4-3(b). By comparison, the degradation in RSNM for the half-selected bits using the negative BL assist, Figure 4-3(b), results in less degradation for the half-selected RSNM. This is an advantage of the negative BL assist over the WL boost. However, as the array supply droops, the negative BL bias can eventually become limited by leakage to the substrate as the forward bias diode begins turning on. To overcome the problematic stability concern for the half-selected bits during a write, a read assist such as VDDc boost may be applied to the non-selected columns. Alternatively, the array architecture can be designed so that the half-select is avoided and all bits on the asserted WL are latched during a write operation.

4.3 Discussion

The characteristic features of the margin/delay plot for read assist were examined. Figure 4-5(a) reveals that the read assist limit contour (ALC) asymptotically approaches the hold SNM (HSNM) limit with increased delay. A simple model is used to describe the observed contour shape. With some simplification, as defined in Appendix 1, the read delay can be approximated by the following relationship:

$$\tau(Vwl, Vddc, Vtn0) = \frac{2 \cdot C_{BL} \cdot Lpg \cdot \Delta V_{BL} \cdot (\psi \cdot \gamma x - \psi \cdot \gamma n - Vtn0 \cdot \beta \cdot \gamma x + Vddc \cdot \beta \cdot \gamma x)}{Wpg \cdot \psi \cdot \beta \cdot kn \cdot \gamma x \cdot (Vtn0 - Vddc) \cdot (2 \cdot Vtn0 - 2 \cdot Vwl + \psi)} \qquad (4\text{-}2)$$

Where ψ represents the velocity saturation value, γn is the body coefficient, γx is a linear approximation factor of the body effect as Vsb increases. Vtn0 is the NMOS threshold voltage with Vsb=0 and β is the PD NMOS W/L divided by the PG NMOS W/L value. The RSNM can be described as linear relationship with bias using the empirically derived sensitivity value obtained in chapter 3. The resulting analytical solution is given in Fig. 4-5(b).

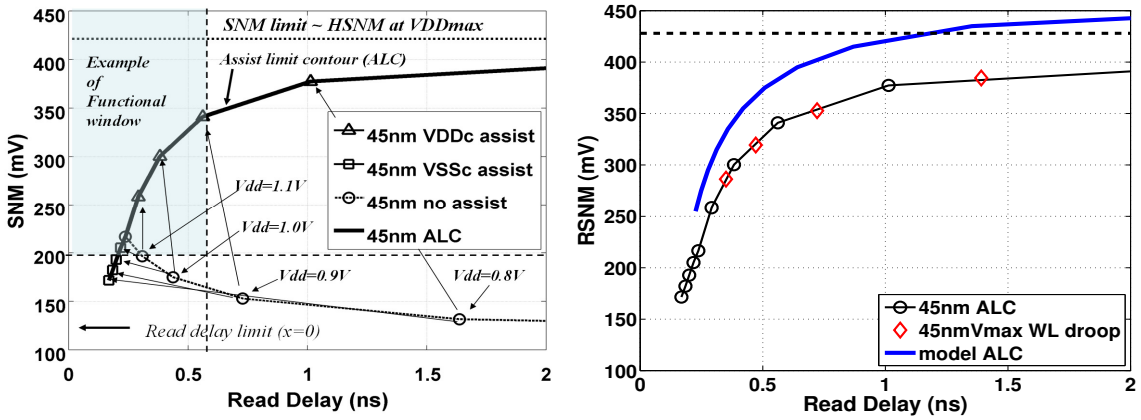


Figure 4-5 (a) The VDDc/VSSc defined read assist limit contour (ALC) as defined by the margin/delay space for 45nm LP PTM 6T SRAM. (b) Analytical ALC model derived using SNM sensitivity with read delay in (4-2).

This general relationship may be anticipated as the latch strength is increased relative to the noise source, the SNM upper limit will approach the HSNM with VDD=Vmax. For the case where the NWELL potential is tied to the array VDD (VDDc), the upper limit will be equal to the hold SNM. This can be more clearly seen from the butterfly curves. Figure 4-6(a) plots the butterfly curves when the VDDc assist method is used with increased assist bias. The characteristic shape evolves with increased assist bias, becoming more similar to the hold SNM shape. Because the performance implications of achieving this limit are in most cases not practical, the more relevant portion of the ALC is across the intersection of the functional window as shown schematically as a shaded region in Figure 4-5(a).

For a given set of technology bias constraints, a contour line defining the upper most noise margin at a given read delay for the technology may be derived. Figure 4-6(b) plots the read ALC mapped across margin/delay space for the LP-PTM technologies from 65nm to 22nm node. Note that for each technology node, the ALC exhibits a similar shape. By applying the specific functional window as determined by the use conditions, array size, and yield requirements, one may follow this approach to assess the viability of the bias based assist methods based on the overlap of the ALC and functional window.

For designs requiring both read and write assist, the yield limiting condition, if not otherwise addressed, may then be the stability upset half-selected bits during a write operation. As shown in Figures 4-3(a) and 4-3(b), the half-selected bits may be adversely affected by the choice of write assist method. This challenge may be addressed by adding an additional bias based assist solution or by array architecture changes which in effect avoid the half-select condition.

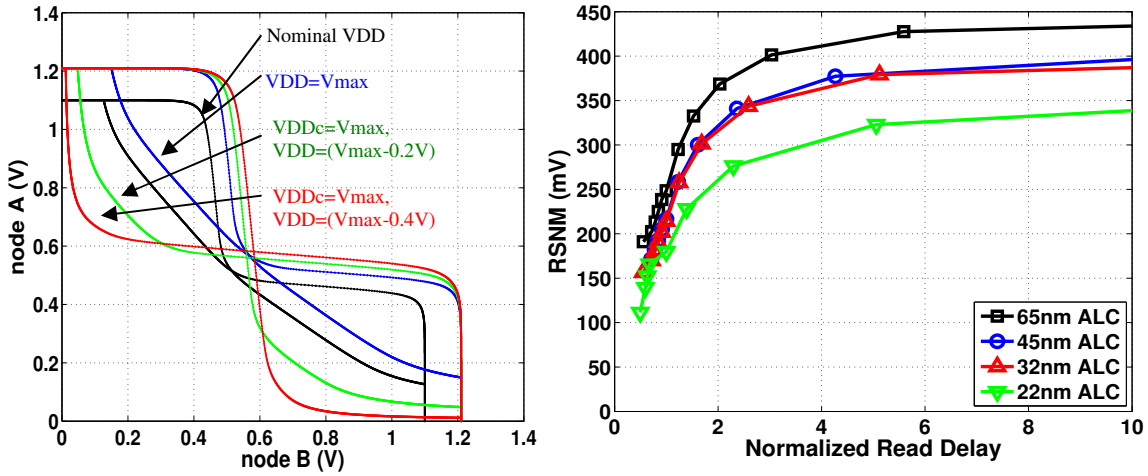


Figure 4-6 (a) Butterfly curves for nominal, Vmax and two VDDc assist cases for the 45nm LP technology. (b) Read assist limit contour (ALC) profiles defined by the VDDc and VSSc terminals for the scaled LP-PTM technologies plotted in the margin/delay space.

4.4 Conclusions

Continued scaling of the planar 6T SRAM will necessitate increased reliance on assist methods to overcome reduced functional yield margins. Because added assist features will incur costs in design complexity, area, and in most cases power, these factors must be balanced against the potential improvement in soft fail limited yield

margin and performance. For bias based assist methods, bias constraints ultimately limit the margin improvements that can be obtained. The applied voltage bias associated with a given assist method must conform to the existing technology bias constraints.

For write assist, in addition to the Vmax constraint, other combined factors also limit the attainable margins. For read assist, by imposing the Vmax constraint a contour is observed in the margin delay space that reflects the relevant attainable limits of a given assist method. The intersection of the ALC with the functional window requirement provides a means to establish bias based assist limitations for a given technology and bit cell. By accounting for these factors, the Vmax constrained read ALC is mapped across four technology generations to gain additional insight into the extent to which assist methods may continue to compensate for the reduced functional margins with continued scaling of the planar 6T SRAM.

Specific Contributions in the field of future and limits of 6T SRAM to include:

- ***A new constraint based approach for assessing the limits of 6T SRAM designs based on the voltage boundary conditions and sensitivities of assist methods***
- ***Based on application of the margin/delay space analysis approach (developed in the preceding chapter) an assist limit contour (ALC) was shown to exist for each technology node.***
- ***The assist limit contour (ALC) is proposed as a fundamental way of addressing the limits of 6T SRAM.***

Publication @ UVA:

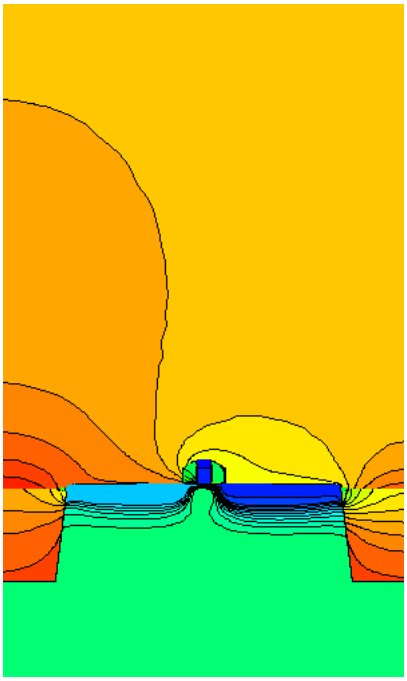
Limits of Bias Based Assist Methods in Nano-Scale 6T SRAM
(Accepted by ISQED for publication in 2010.)

CHAPTER 5

Summary and Future Work

-Abstract

The important role of the 6T SRAM cell in the VLSI industry and the extent to which it can be extended through continued scaling is of enormous economic importance. This body of work provides insight into the challenging and complex cell design constraints being faced by the industry leaders in CMOS process technology today. Understanding these limitations and the mechanisms that underlie them provides insight to identify more optimum circuit solutions. At the time of this work, the worlds leading advanced silicon providers are developing what will be known as the 32nm (or 28nm/32nm) generation technologies. Although the planar 6T SRAM will be, as in previous generations, a dominant factor in the development of this generation of technologies, incorporation of the 8T will become more common for L1 cache applications. Beyond 32nm, it is less clear if the planar 6T cell will again prove to be the dominant memory in large L2 cache and mobile computing applications. This will depend on the success of circuit assist methods in conjunction with emerging technology options.



Subject	Contribution Description	Status	Completion
Variability in nano-scale CMOS	***infrastructure circuit designs for array leakage, bfly curve analysis and Vt mismatch***	Tape out complete in 180nm FDSOI	5/2009
	Design and measurement and analysis of FDSOI device mismatch	-Tape out complete -hardware target 1/2010	**3/2010 target**
Nano-scale CMOS technology	***High Performance and Low Power Transistors Integrated in 65nm Bulk CMOS Technology***	Published with 26 citations	2004
	Development of Low Power PTM Models (90-22nm)	used in two publications	8/2008
	Lateral Ion Implant Straggle and Mask Proximity Effect	Published with 25 citations	2003
Nano-scale SRAM cell design	***SRAM cell design for stability methodology***	Published with 16 citations	2005
	cell topology and design trade-offs for nano-scale CMOS	Develop TCAD tool experience	**5/2010 target**
	Diffusion, implant scattering and shadowing implications on future SRAM cell design	Supported by simulations with TCAD tools	**6/2010 target**
	SRAM cell design beyond 32nm	Concept for publication	**7/2010 target**
Low Power SRAM Technology	***Ultralow-Power SRAM technology***	Published with 31 citations	2003
	Noise Margin and Leakage in Ultra-Low Leakage SRAM Cell Design	Published with 19 citations	2002
Coping with Variability: SRAM circuit assist methods	***Methodology for Evaluation of 6T SRAM Assist Methods in Scaled CMOS Technologies***	Submit for publication	**1/2010 target**
	SRAM voltage control for improved operational margins	Patent 7,313,032	12/2007
Limits of 6T SRAM	***Limits of Bias Based Assist Methods in Nano-Scale 6T SRAM***	Accepted for publication by ISQED 3/2010	11/19/2009
Thesis	***Dissertation manuscript***		**8/2010 target**

APPENDIX 1

Ipd is in the linear mode, therefore:

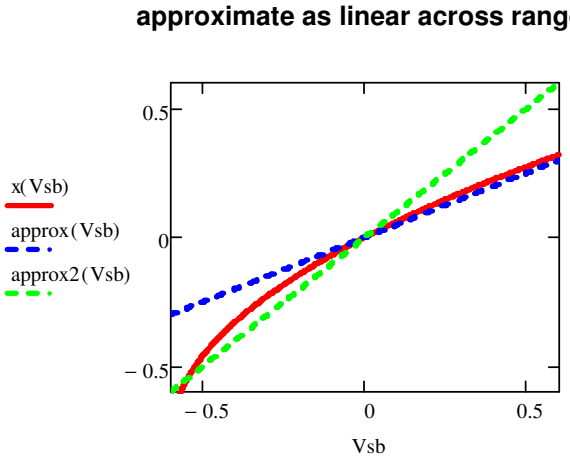
$$Ipd(Va, Vb) = kn \cdot \frac{Wpd}{Lpd} \cdot Vb \cdot \left(Va - Vtn0 - \frac{Vb}{2} \right)$$

Where kn is the product of mobility and Cox, Va is the gate voltage supplied by the latch cross couple. Vb is the internal voltage determined by the voltage divider relationship between the PG and PD NMOS devices.
Ipg in velocity saturation:

$$Ipg(Vb) = kn \cdot \frac{Wpg}{Lpg} \cdot \Psi \cdot \left(Vdd - Vb - Vtn0 + \frac{\gamma n \cdot Vb}{2} - \frac{\Psi}{2} \right) \cdot [1 + \lambda n \cdot (Vdd - Vb)]$$

Ψ is the Velocity saturation voltage

Simplification 1: neglect channel length modulation
Simplification 2: linearize body effect in the following way:



Simplification 3: neglect small Vb/2 term in linear model

This allows the voltage on node b (Vnb) to be expressed as:

$$Vnb(Vwl, Vddc, Vtn0) = \frac{\psi \cdot Vwl - \psi \cdot Vtn0 - \frac{\psi^2}{2}}{\beta \cdot Vddc - \beta \cdot Vtn0 + \psi - \frac{\psi \cdot \gamma n}{\gamma x}}$$

By substitution the read current can be written as:

$$Ir(Vwl, Vddc, Vtn0) = kn \cdot \frac{Wpg}{Lpg} \cdot \psi \cdot \left(Vwl - Vnb(Vwl, Vddc, Vtn0) - Vtn0 + \frac{\gamma n \cdot Vnb(Vwl, Vddc, Vtn0)}{\gamma x} - \frac{\psi}{2} \right)$$

The read delay is expressed as:

$$\tau_r(Vwl, Vddc, Vtn0) = \frac{C_{BL} \cdot \Delta V_{BL}}{Ir(Vwl, Vddc, Vtn0)}$$

Where C_{BL} is the BL capacitance and ΔV_{BL} is the delta voltage that must be developed on the BL to successfully read. 100mV is used in this calculation.

The read delay as s function of Vtn, Vwl and Vddc is then expressed as:

$$\tau(Vwl, Vddc, Vtn0) = \frac{2 \cdot C_{BL} \cdot Lpg \cdot \Delta V_{BL} \cdot (\psi \cdot \gamma x - \psi \cdot \gamma n - Vtn0 \cdot \beta \cdot \gamma x + Vddc \cdot \beta \cdot \gamma x)}{Wpg \cdot \psi \cdot \beta \cdot kn \cdot \gamma x \cdot (Vtn0 - Vddc) \cdot (2 \cdot Vtn0 - 2 \cdot Vwl + \psi)}$$

References

[1] C. C. Wu, Y. K. Leung, C. S. Chang, M. H. Tsai, H. T. Huang, D. W. Lin, Y. M. Sheu, C. H. Hsieh, W. J. Liang, L. K. Han, W. M. Chen, S. Z. Chang, S. Y. Wu, S. S. Lin, H. C. Lin, C. H. Wang, P. W. Wang, T. L. Lee, C. Y. Fu, C. W. Chang, S. C. Chen, S. M. Jang, S. L. Shue, H. T. Lin, Y. C. See, Y. J. Mii, C. H. Diaz, B. J. Lin, M. S. Liang, and Y. C. Sun, "A 90-nm cmos device technology with high-speed, general-purpose, and low-leakage transistors for system on chip applications," in *Proc. Digest. International Electron Devices Meeting IEDM '02*, 8–11 Dec. 2002, pp. 65–68.

[2] Z. Luo, A. Steegen, M. Eller, R. Mann, C. Baiocco, P. Nguyen, L. Kim, M. Hoinkis, V. Ku, V. Klee, F. Jamin, P. Wrschka, P. Shafer, W. Lin, S. Fang, A. Ajmera, W. Tan, D. Park, R. Mo, J. Lian, D. Vietzke, C. Coppock, A. Vayshenker, T. Hook, V. Chan, K. Kim, A. Cowley, S. Kim, E. Kaltalioglu, B. Zhang, S. Marokkey, Y. Lin, K. Lee, H. Zhu, M. Weybright, R. Rengarajan, J. Ku, T. Schiml, J. Sudijono, I. Yang, and C. Wann, "High performance and low power transistors integrated in 65nm bulk cmos technology," in *Proc. IEDM Technical Digest Electron Devices Meeting IEEE International*, 2004, pp. 661–664.

[3] Z. Luo, N. Rovedo, S. Ong, B. Phoong, M. Eller, H. Utomo, C. Ryou, H. Wang, R. Stierstorfer, L. Clevenger, S. Kim, J. Toomey, D. Sciacca, J. Li, W. Wille, L. Zhao, L. Teo, T. Dyer, S. Fang, J. Yan, O. Kwon, D. Park, J. Holt, J. Han, V. Chan, T. K. J. Yuan, H. Lee, S. Lee, A. Vayshenker, Z. Yang, C. Tian, H. Ng, H. Shang, M. Hierlemann, J. Ku, J. Sudijono, and M. Ieong, "High performance transistors featured in an aggressively scaled 45nm bulk cmos technology," in *Proc. IEEE Symposium on VLSI Technology*, 12–14 June 2007, pp. 16–17.

[4] S.-Y. Wu, C. W. Chou, C. Y. Lin, M. C. Chiang, C. K. Yang, M. Y. Liu, L. C. Hu, C. H. Chang, P. H. Wu, H. F. Chen, S. Y. Chang, S. H. Wang, P. Y. Tong, Y. L. Hsieh, J. J. Liaw, K. H. Pan, C. H. Hsieh, C. H. Chen, J. Y. Cheng, C. H. Yao, W. K. Wan, T. L. Lee, K. T. Huang, K. C. Lin, L. Y. Yeh, K. C. Ku, S. C. Chen, H. J. Lin, S. M. Jang, Y. C. Lu, J. H. Shieh, M. H. Tsai, J. Y. Song, K. S. Chen, V. Chang, S. M. Cheng, S. H. Yang, C. H. Diaz, Y. C. See, and M. S. Liang, "A 32nm cmos low power soc platform technology for foundry applications with functional high density sram," in *Proc. IEEE International Electron Devices Meeting IEDM 2007*, 10–12 Dec. 2007, pp. 263–266.

[5] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of mos transistors," vol. 24, no. 5, pp. 1433–1439, Oct 1989.

[6] J. Mc Ginley, O. Noblanc, C. Julien, S. Parihar, K. Rochereau, R. Difrenza, and P. Llinares, "Impact of pocket implant on mosfet mismatch for advanced cmos technology," in *Proc. International Conference on Microelectronic Test Structures ICMTS '04*, 22–25 March 2004, pp. 123–126.

[7] H. S. Yang, R. Wong, R. Hasumi, Y. Gao, N. S. Kim, D. H. Lee, S. Badrudduza, D. Nair, M. Ostermayr, H. Kang, H. Zhuang, J. Li, L. Kang, X. Chen, A. Thean, F. Arnaud, L. Zhuang, C. Schiller, D. P. Sun, Y. W. Teh, J. Wallner, Y. Takasu, K. Stein, S. Samavedam, D. Jaeger, C. V. Baiocco, M. Sherony, M. Khare, C. Lage, J. Pape, J. Sudijono, A. L. Steegen, and S. Stiffler, "Scaling of 32nm low power SRAM with

high-K metal gate," in *Proc. IEEE International Electron Devices Meeting IEDM 2008*, 15–17 Dec. 2008, pp. 1–4.
[8] A. Asenov and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1-& mosfet's with epitaxial and &doped channels," vol. 46, no. 8, pp. 1718–1724, Aug. 1999.
[9] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of mos sram cells," vol. 22, no. 5, pp. 748–754, Oct 1987.
[10] T. N. Y. Taur, *Fundamentals of Modern VLSI Devices*, U. Cambridge, Ed. Cambridge Univ. Press, 1998.
[11] R. Mann, W. Abadeer, M. Breitwisch, O. Bula, J. Brown, B. Colwill, P. Cottrell, W. Crocco, S. Furkay, M. Hauser, T. Hook, D. Hoyniak, J. Johnson, C. Lam, R. Mih, J. Rivard, A. Moriwaki, E. Phipps, C. Putnam, B. Rainey, J. Toomey, and M. Younus, "Ultralow-power sram technology," *IBM J. Res. & Dev.*, vol. Vol. 47, pp. 553–566, 2003.
[12] R. Baumann, "The impact of technology scaling on soft error rate performance and limits to the efficacy of error correction," in *Proc. Digest. International Electron Devices Meeting IEDM '02*, 8–11 Dec. 2002, pp. 329–332.
[13] P. E. Dodd and L. W. Massengill, "Basic mechanisms and modeling of single-event upset in digital microelectronics," vol. 50, no. 3, pp. 583–602, June 2003.
[14] N. Seifert, P. Slankard, M. Kirsch, B. Narasimham, V. Zia, C. Brookreson, A. Vo, S. Mitra, B. Gill, and J. Maiz, "Radiation-induced soft error rates of advanced cmos bulk devices," in *Proc. 44th Annual. IEEE International Reliability Physics Symposium*, 26–30 March 2006, pp. 217–225.
[15] P. Hazucha and C. Svensson, "Impact of cmos technology scaling on the atmospheric neutron soft error rate," vol. 47, no. 6, pp. 2586–2594, Dec. 2000.
[16] K. Kang, H. Kufluoglu, K. Roy, and M. Ashraful Alam, "Impact of negative-bias temperature instability in nanoscale sram array: Modeling and analysis," vol. 26, no. 10, pp. 1770–1781, Oct. 2007.
[17] A. T. Krishnan, V. Reddy, D. Aldrich, J. Raval, K. Christensen, J. Rosal, C. O'Brien, R. Khamankar, A. Marshall, W. K. Loh, R. McKee, and S. Krishnan, "Sram cell static noise margin and vmin sensitivity to transistor degradation," in *Proc. International Electron Devices Meeting IEDM '06*, 11–13 Dec. 2006, pp. 1–4.
[18] S. Nalam and B. H. Calhoun, "Asymmetric sizing in a 45nm 5t sram to improve read stability over 6t," in *Proc. IEEE Custom Integrated Circuits Conference CICC '09*, 13–16 Sept. 2009, pp. 709–712.
[19] R. E. Aly and M. A. Bayoumi, "Low-power cache design using 7t sram cell," vol. 54, no. 4, pp. 318–322, April 2007.
[20] S. Lin, Y.-B. Kim, and F. Lombardi, "A 32nm sram design for low power and high stability," in *Proc. 51st Midwest Symposium on Circuits and Systems MWSCAS 2008*, 10–13 Aug. 2008, pp. 422–425.
[21] Z. Liu and V. Kursun, "Characterization of a novel nine-transistor sram cell," vol. 16, no. 4, pp. 488–492, April 2008.
[22] B. H. Calhoun and A. Chandrakasan, "A 256kb sub-threshold sram in 65nm cmos," in *Proc. Digest of Technical Papers. IEEE International Solid-State Circuits Conference ISSCC 2006*, 6–9 Feb. 2006, pp. 2592–2601.
[23] S. Okumura, Y. Iguchi, S. Yoshimoto, H. Fujiwara, H. Noguchi, K. Nii, H. Kawaguchi, and M. Yoshimoto, "A 0.56-v 128kb 10t sram using column line assist

(cla) scheme," in *Proc. Quality of Electronic Design Quality of Electronic Design ISQED 2009*, 16–18 March 2009, pp. 659–663.
[24] H. Noguchi, S. Okumura, Y. Iguchi, H. Fujiwara, Y. Morita, K. Nii, H. Kawaguchi, and M. Yoshimoto, "Which is the best dual-port sram in 45-nm process technology? &8t, 10t single end, and 10t differential &," in *Proc. IEEE International Conference on Integrated Circuit Design and Technology and Tutorial ICICDT 2008*, 2–4 June 2008, pp. 55–58.
[25] L. Chang, D. M. Fried, J. Hergenrother, J. W. Sleight, R. H. Dennard, R. K. Montoye, L. Sekaric, S. J. McNab, A. W. Topol, C. D. Adams, K. W. Guarini, and W. Haensch, "Stable SRAM cell design for the 32 nm node and beyond," in *Proc. Digest of Technical Papers VLSI Technology 2005 Symposium on*, 14–16 June 2005, pp. 128–129.
[26] L. Chang, R. K. Montoye, Y. Nakamura, K. A. Batson, R. J. Eickemeyer, R. H. Dennard, W. Haensch, and D. Jamsek, "An 8t-sram for variability tolerance and low-voltage operation in high-performance caches," vol. 43, no. 4, pp. 956–963, April 2008.
[27] M. Ishida, T. Kawakami, A. Tsuji, N. Kawamoto, M. Motoyoshi, and N. Ouchi, "A novel 6t-sram cell technology designed with rectangular patterns scalable beyond 0.18 & generation and desirable for ultra high speed operation," in *Proc. International Electron Devices Meeting IEDM '98 Technical Digest*, 6–9 Dec. 1998, pp. 201–204.
[28] S. Thompson, M. Alavi, R. Arghavani, A. Brand, R. Bigwood, J. Brandenburg, B. Crew, V. Dubin, M. Hussein, P. Jacob, C. Kenyon, E. Lee, B. Mcintyre, Z. Ma, P. Moon, P. Nguyen, M. Prince, R. Schweinfurth, S. Sivakumar, P. Smith, M. Stettler, S. Tyagi, M. Wei, J. Xu, S. Yang, and M. Bohr, "An enhanced 130 nm generation logic technology featuring 60 nm transistors optimized for high performance and low power at 0.7 - 1.4 v," in *Proc. IEDM Technical Digest Electron Devices Meeting International*, 2–5 Dec. 2001, pp. 11.6.1–11.6.4.
[29] P. Bai, C. Auth, S. Balakrishnan, M. Bost, R. Brain, V. Chikarmane, R. Heussner, M. Hussein, J. Hwang, D. Ingerly, R. James, J. Jeong, C. Kenyon, E. Lee, S. H. Lee, N. Lindert, M. Liu, Z. Ma, T. Marieb, A. Murthy, R. Nagisetty, S. Natarajan, J. Neirynck, A. Ott, C. Parker, J. Sebastian, R. Shaheed, S. Sivakumar, J. Steigerwald, S. Tyagi, C. Weber, B. Woolery, A. Yeoh, K. Zhang, and M. Bohr, "A 65nm logic technology featuring 35nm gate lengths, enhanced channel strain, 8 cu interconnect layers, low-k ild and 0.57 &sram cell," in *Proc. IEDM Technical Digest Electron Devices Meeting IEEE International*, 13–15 Dec. 2004, pp. 657–660.
[30] T. B. Hook, J. Brown, P. Cottrell, E. Adler, D. Hoyniak, J. Johnson, and R. Mann, "Lateral ion implant straggle and mask proximity effect," vol. 50, no. 9, pp. 1946–1951, Sept. 2003.
[31] K.-L. Cheng, C. C. Wu, Y. P. Wang, D. W. Lin, C. M. Chu, Y. Y. Tarng, S. Y. Lu, S. J. Yang, M. H. Hsieh, C. M. Liu, S. P. Fu, J. H. Chen, C. T. Lin, W. Y. Lien, H. Y. Huang, P. W. Wang, H. H. Lin, D. Y. Lee, M. J. Huang, C. F. Nieh, L. T. Lin, C. C. Chen, W. Chang, Y. H. Chiu, M. Y. Wang, C. H. Yeh, F. C. Chen, Y. H. Chang, S. C. Wang, H. C. Hsieh, M. D. Lei, K. Goto, H. J. Tao, M. Cao, H. C. Tuan, C. H. Diaz, Y. J. Mii, and C. M. Wu, "A highly scaled, high performance 45 nm bulk logic cmos technology with 0.242 &sram cell," in *Proc. IEEE International Electron Devices Meeting IEDM 2007*, 10–12 Dec. 2007, pp. 243–246.

[32] J. B. Johnson, T. B. Hook, and Y.-M. Lee, "Analysis and modeling of threshold voltage mismatch for cmos at 65 nm and beyond," vol. 29, no. 7, pp. 802–804, July 2008.
[33] S. Bordez, A. Cathignol, and K. Rochereau, "A continuous model for mosfet vt matching considering additional length effects," in *Proc. IEEE International Conference on Microelectronic Test Structures ICMTS '07*, 19–22 March 2007, pp. 226–229.
[34] S. Ohbayashi, M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Imaoka, Y. Oda, T. Yoshihara, M. Igarashi, M. Takeuchi, H. Kawashima, Y. Yamaguchi, K. Tsukamoto, M. Inuishi, H. Makino, K. Ishibashi, and H. Shinohara, "A 65-nm SoC embedded 6T-SRAM designed for manufacturability with read and write operation stabilizing circuits," vol. 42, no. 4, pp. 820–829, April 2007.
[35] O. Hirabayashi, A. Kawasumi, A. Suzuki, Y. Takeyama, K. Kushida, T. Sasaki, A. Katayama, G. Fukano, Y. Fujimura, T. Nakazato, Y. Shizuki, N. Kushiyama, and T. Yabe, "A process-variation-tolerant dual-power-supply SRAM with 0.179μm2 cell in 40nm CMOS using level-programmable wordline driver," in *Proc. IEEE International Solid-State Circuits Conference - Digest of Technical Papers ISSCC 2009*, 8–12 Feb. 2009, pp. 458–459,459a.
[36] D. P. Wang, H. J. Liao, H. Yamauchi, Y. H. Chen, Y. L. Lin, S. H. Lin, D. C. Liu, H. C. Chang, and W. Hwang, "A 45nm dual-port SRAM with write and read capability enhancement at low voltage," in *Proc. IEEE International SOC Conference*, 26–29 Sept. 2007, pp. 211–214.
[37] B. Mohammad, M. Saint-Laurent, P. Bassett, and J. Abraham, "Cache design for low power and high yield," in *Proc. 9th International Symposium on Quality Electronic Design ISQED 2008*, 17–19 March 2008, pp. 103–107.
[38] K. Nii, M. Yabuuchi, Y. Tsukamoto, S. Ohbayashi, Y. Oda, K. Usui, T. Kawamura, N. Tsuboi, T. Iwasaki, K. Hashimoto, H. Makino, and H. Shinohara, "A 45-nm single-port and dual-port SRAM family with robust read/write stabilizing circuitry under DVFS environment," in *Proc. IEEE Symposium on VLSI Circuits*, 18–20 June 2008, pp. 212–213.
[39] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Reduction of parametric failures in sub-100-nm SRAM array using body bias," vol. 27, no. 1, pp. 174–183, Jan. 2008.
[40] M. Yamaoka, N. Maeda, Y. Shinozaki, Y. Shimazaki, K. Nii, S. Shimada, K. Yanagisawa, and T. Kawahara, "Low-power embedded SRAM modules with expanded margins for writing," in *Proc. Digest of Technical Papers Solid-State Circuits Conference ISSCC. 2005 IEEE International*, 10–10 Feb. 2005, pp. 480–611.
[41] M. Khellah, Y. Ye, N. S. Kim, D. Somasekhar, G. Pandya, A. Farhang, K. Zhang, C. Webb, and V. De, "Wordline & bitline pulsing schemes for improving SRAM cell stability in low-Vcc 65nm CMOS designs," in *Proc. Digest of Technical Papers VLSI Circuits 2006 Symposium on*, 2006, pp. 9–10.
[42] A. Bhavnagarwala, S. Kosonocky, C. Radens, K. Stawiasz, R. Mann, Q. Ye, and K. Chin, "Fluctuation limits & scaling opportunities for CMOS SRAM cells," in *Proc. IEDM Technical Digest Electron Devices Meeting IEEE International*, 5–5 Dec. 2005, pp. 659–662.
[43] A. J. Bhavnagarwala, S. Kosonocky, C. Radens, Y. Chan, K. Stawiasz, U. Srinivasan, S. P. Kowalczyk, and M. M. Ziegler, "A sub-600-mv, fluctuation tolerant 65-nm CMOS SRAM array with dynamic cell biasing," vol. 43, no. 4, pp. 946–955, April 2008.

[44] M. H. Abu-Rahma, M. Anis, and S. S. Yoon, "A robust single supply voltage SRAM read assist technique using selective precharge," in *Proc. 34th European Solid-State Circuits Conference ESSCIRC 2008*, 15–19 Sept. 2008, pp. 234–237.
[45] Y. H. Chen, W. M. Chan, S. Y. Chou, H. J. Liao, H. Y. Pan, J. J. Wu, C. H. Lee, S. M. Yang, Y. C. Liu, and H. Yamauchi, "A 0.6V 45nm adaptive dual-rail SRAM compiler circuit design for lower VDDmin VLSIs," in *Proc. IEEE Symposium on VLSI Circuits*, 18–20 June 2008, pp. 210–211.
[46] Y. Chung and S.-H. Song, "Implementation of low-voltage static RAM with enhance data stability and circuit speed," *Microelectronics Journal*, vol. 40, pp. 944–951, 2009.
[47] K. Zhang, U. Bhattacharya, Z. Chen, F. Hamzaoglu, D. Murray, N. Vallepalli, Y. Wang, B. Zheng, and M. Bohr, "A 3-GHz 70-Mb SRAM in 65-nm CMOS technology with integrated column-based dynamic power supply," vol. 41, no. 1, pp. 146–151, Jan. 2006.
[48] M. Yamaoka, K. Osada, and K. Ishibashi, "0.4-V logic-library-friendly SRAM array using rectangular-diffusion cell and delta-boosted-array voltage scheme," vol. 39, no. 6, pp. 934–940, June 2004.
[49] M. Iijima, K. Seto, M. Numa, A. Tada, and T. Ipposhi, "Low power SRAM with boost driver generating pulsed word line voltage for sub-1V operation," *JCP 3*, vol. 5, pp. 34–40, 2008.
[50] N. Shibata, H. Kiya, S. Kurita, H. Okamoto, M. Tan'no, and T. Douseki, "A 0.5-V 25-MHz 1-mW 256-kb MTCMOS/SOI SRAM for solar-power-operated portable personal digital equipment - sure write operation by using step-down negatively overdriven bitline scheme," vol. 41, no. 3, pp. 728–742, March 2006.
[51] H. Pilo, J. Barwin, G. Braceras, C. Browning, S. Burns, J. Gabric, S. Lamphier, M. Miller, A. Roberts, and F. Towler, "An SRAM design in 65nm and 45nm technology nodes featuring read and write-assist circuits to expand operating voltage," in *Proc. Digest of Technical Papers VLSI Circuits 2006 Symposium on*, 2006, pp. 15–16.
[52] T. Suzuki, H. Yamauchi, Y. Yamagami, K. Satomi, and H. Akamatsu, "A stable 2-port SRAM cell design against simultaneously read/write-disturbed accesses," vol. 43, no. 9, pp. 2109–2119, Sept. 2008.
[53] M. Yamaoka, N. Maeda, Y. Shimazaki, and K. Osada, "65nm low-power high-density SRAM operable at 1.0V under 3σ systematic variation using separate Vth monitoring and body bias for NMOS and PMOS," in *Proc. Digest of Technical Papers. IEEE International Solid-State Circuits Conference ISSCC 2008*, 3–7 Feb. 2008, pp. 384–622.